5-15-2007

Essays on the Value of a Statistical Life

Ikuho Kochi

PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this dissertation may be granted by the author or, in his or her absence, the professor under whose direction it was written or, in his or her absence, by the Dean of the Andrew Young School of Policy Studies. Such quoting, copying, or publishing must be solely for scholarly purposes and must not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential gain will not be allowed without written permission of the author.

____

Signature of the Author

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Ikuho Kochi
4470 South Lemay Ave Apt 403
Fort Collins, CO 80525

The director of this dissertation is:

Laura O. Taylor
Department of Economics
Andrew Young School of Policy Studies
Georgia State University

Users of this dissertation not regularly enrolled as students at Georgia State University are required to attest acceptance of the preceding stipulations by signing below. Libraries borrowing this dissertation for the use of their patrons are required to see that each user records here the information requested.

Name of User	Address	Date	Type of use (Examination only or copying)

ESSAYS ON THE VALUE OF A STATISTICAL LIFE

BY

IKUHO KOCHI

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree
of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2007

ACCEPTANCE

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair: Laura O. Taylor

Committee: H. Spencer Banzhaf
Susan K. Laury
Kenneth E. McConnell
Mary Beth Walker

Electronic Version Approved:
Roy W. Bahl, Dean
Andrew Young School of Policy Studies
Georgia State University
May 2007

ACKNOWLEDGEMENTS

First, I would like to thank my dissertation chair, Dr. Laura O. Taylor, for her guidance, and extensive support and advice to complete this dissertation. She is my role model as a researcher, teacher, mother, wife, and dissertation chair. I also would like to express my appreciation for my dissertation committee members, Dr. Mary Beth Walker, Dr. H. Spencer Banzhaf, Dr. Susan Laury and Dr. Kenneth E. McConnell for their insightful and valuable comments.

I am also thankful for the people and organizations that enabled me to complete this dissertation. I thank Scott Richardson at the BLS for allowing access to a special research file. I also appreciate Samuel Meyer, Jason Shippy, Maury Gittleman at the BLS for their dedicated and efficient work. I am grateful for Dr. Carol Scotton for providing me her data. I also thank Georgia State University and Dr. Laura O. Taylor for funding my dissertation work. I am very thankful for Dr. Kenneth E. McConnell and his wife Dr. Virginia D. McConnell, and my good friend Mercy Mvundura and her family for providing me accommodations during my research trips in Washington, D.C.

I thank all my friends in Atlanta for their support during this hard doctoral student life. Without laughter with you, there was no way that I could survive the last five years. I especially dedicate my thanks to Djesika Amendah, Costanza Meneghetti, Cristian Sepulveda, and Project Ladies members for their constant friendship. Of course, my deepest appreciation goes to Dr. Raul Alberto Ponce Rodriguez. You are a gift from God to enlighten my life.

Lastly, I would like to dedicate this dissertation to my mother, Mrs. Takako Kochi. Without her belief in me over years, I would not be writing acknowledgements for my dissertation right now.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS vi

List of Tables x

List of Figures xiii

List of Appendices xiv

Abstract xv

Chapter I Introduction 1

Chapter II Theory of Compensating Wage Differentials 9

Preferences over Risks and Wages: the Worker 10

Firm's Production and Risk 13

Underlying Assumptions 16

Extensions Examined in This Dissertation 19

Chapter III Data 25

Data for Panel Study (Chapter 4) 27

Labor force data. 27

Occupational fatal risk data. 33

Data for Chapter 5 41

Labor force data. 41

Local area unemployment statistics. 44

Economic activity in each MSA. 44

The vehicle miles traveled (VMT). 45

Occupational fatal incidences. 45

Area specific employment level. 48

Fatal risk rate. 50

Non-fatal injury incidence. 52

Fatal and non-fatal risks in the sample. 55

Chapter IV   Panel Data Analysis of Hedonic Wage Model..78

Endogeneity in Hedonic Wage Models 78

Basic FD and FE model..82

Endogeneity in Panel Models 83

Results 85

Sensitivity Analysis..92

Conclusions...96

Chapter V   Transferability of VSL: Hedonic Wage Analysis 107

Literature Review...110

Empirical Model 112

Results 114

Conclusions...124

Chapter VI   Conclusions...136

References…………………………………………………………………………156

Vita………………………………………………………………………………...161

LIST OF TABLES

Table		Page
1	Definition of Variables and Summary Statistics	58
2	The Number of Observations Changing Jobs between Major Occupation Group	60
3	The Number of Observations Changing Jobs between Major Industry Group	61
4	The Risk Rate Change in Each Major Occupation Group	62
5	The Risk Rate Change in Each Major Industry Group	63
6	The Summary of Risk Changes in Different Groups of Workers	64
7	The Summary of Risk Changes in Different Groups of Workers Who Changes Demographic Status	65
8	Variable Definition, Source and Summary statistics	66
9	Summary of Fatal Incidences by Occupation and Event (1992-2002 CFOI)	67
10	MSAs with Top Five Number of Deaths by Occupation and Event (1992-2002 CFOI)	68
11	State with Top Five Number of Death Incidences by Occupation and Event (1992-2002 CFOI)	69
12	National Trend of Employment Growth in Driving Occupation	70
13	Summary of Fatal Incidence by Occupation and Event in the 1998-2002 CFOI	70
14	Mean and Standard Deviation of Fatal Risk Rate by Occupation and Event (1992-2002 CFOI)	71
15	Mean and Standard Deviation of MSA Level Fatal Risk Rate by Occupation and Event (1992-2002 CFOI) After Dropping MSAs With Less Than 100 Employees	71

16 Mean and Standard Deviation of Fatal Risk Rate by Occupation and Event Based on 1998-2002 CFOI 72

17 Mean and Standard Deviation of MSA Level Fatal Risk Rate by Occupation and Event After Dropping MSAs with Less Than 100 Employees Based on 1998-2002 CFOI 72

18 Missing Injury Data by State 73

19 State Level Non-fatal Injury Rates by Occupation Based on the IFF 1998-2002 74

20 State Level Non-fatal Injury Rates by Occupation Based on the IFF 1998-2002 74

21 Summary of *MSA* Level Fatal Risk Rates for the Sample of Workers (Risk Creased Based on 1992-2002 CFOI) 75

22 Summary of *State* Level Fatal Risk Rates for the Sample of Workers (Risk Creased Based on 1992-2002 CFOI) 75

23 Summary of *MSA* Level Fatal Risk Rates for the Sample of Workers (Risk Creased Based on 1998-2002 CFOI) 76

24 Summary of *State* Level Fatal Risk Rates for the Sample of Workers (Injury Risk Creased Based on 1992-2002 IFF) 76

25 Summary of State Level Injury Rates for the Sample of Workers (Risk Creased Based on 1992-2002 CFOI) 77

26 Summary of State Level Injury Rates for the Sample of Workers (Risk Creased Based on 1998-2002 IFF) 77

27 Cross-section, Fixed Effect and First Difference Regression Results 100

28 First Stage Regression Results With Second Set of Instrumental Variables 101

29 Second Stage (IV-FE and IV-FD) Regression Results 102

30 IV-FE and IV-FD Results With First Sets of Instruments 103

31 Summary Statistics of Key Characteristics of Job Changers and Non-job Changer Sample 104

32 Cross-section, Fixed Effect and First Difference Regression Results With Only Job Changer Sample 105

33 IV-FE and IV-FD Results With Second Type of Instruments: Job-changer Sample 106

34 Summary of Estimated Coefficients 106

35 Regression Results 129

36 Results With MSA Level Violent Risk and S*tate* Level Non-violent Risk Without MSA<100 130

37 Results with MSA Level or *State* Level Fatal Risks Without MSA<100 131

38 Results With Allowing Correlation Within a Geographic Unit 132

39 Different Geographic Risk Level by Type of Driver With Allowing Correlation Within a Geographic Unit 133

40 Results with Limited Risk Range (With State Dummy and Allowing Correlation Within MSA) 134

41 Results With MSA Variables With Allowing Correlation Within MSA 135

LIST OF FIGURES

Figure Page

1 HW equilibrium schedule for heterogeneous workers and heterogeneous firms………………………………………………………..22

2 HW equilibrium schedule with workers who underestimate risk levels.............23

3 HW equilibrium schedules for workers with different productivity...................24

LIST OF APPENDICES

Appendix A CFOI Research File Elements 145

Appendix B Industry Group Classification...146

Appendix C Occupation Group Classification 147

Appendix D Occupation Group Mapping...148

Appendix E Industry Group Mapping 149

Appendix F Definition of Occupational Drivers 153

Appendix G Complete Set of Results With Allowing Correlation Within a Geographic Unit...155

ABSTRACT

ESSAYS ON THE VALUE OF STATISTICAL LIFE

By

IKUHO KOCHI

May 2007

Committee Chair: Dr. Laura O. Taylor

Major Department: Economics

This dissertation addresses two important issues in the literature estimating the Value of a Statistical Life. The first issue is the potential endogeneity bias in cross-section hedonic wage models. The second issue is the transferability of the VSL between different policy contexts.

To address the first issue, we estimate cross-section and panel hedonic wage models to identify the bias due to the time-invariant worker heterogeneity. We also consider potential endogeneity bias due to measurement error associated with risk variable, time-variant worker heterogeneity and simultaneity between wage and risk in panel models. We obtain labor market data from the 1996 Survey of Income and Program Participation panel, and occupational fatal risk data from Scotton (2000). We find that the cross-section hedonic wage model is significantly biased upward due to unobserved time-invariant worker heterogeneity, but not from time-variant worker heterogeneity or simultaneity between wage and risk. Our results are sensitive to the inclusion of industry variables, but not sensitive to the sample of workers used in estimation.

To address the second issue, we examine whether or not workers and firms differentiate heterogeneous risks to determine the risk-wage compensation levels. We focus on two very different fatal risks in terms of the degree of workers' control over the risk and the degree of dread associated with risk: violent assaults and risks related to non-violent events. We use occupational drivers to mitigate potential unobserved heterogeneity of job characteristics and measurement error associated with risk variables. The labor market data comes from the basic CPS, and the occupation-geographic specific risk rates for each cause of death are created from the non-public Census of Fatal Occupational Injuries. We find that occupational drivers require larger compensation to accept a marginal increase of violent risk as compared to non-violent risk. This is true for both fatal and non-fatal risks. Our results are quite robust. This study suggests that current direct use of VSL obtained from hedonic wage studies in benefit estimation of various governmental programs should be reconsidered.

Chapter I

Introduction

The Value of a Statistical Life (VSL) is an important component of benefit estimates for many governmental programs that intend to reduce premature deaths. The VSL is a society's aggregated willingness to pay to save one anonymous person's life (Fisher, Chestnut, & Violette, 1989). Due to the difficulties of estimating the value of life directly, the VSL is estimated based on individual behavior as related to risk-dollar tradeoffs. For example, in a society of 100,000 people, if each individual is willing to pay $10 to reduce the risk of death from 2 in 100,000 to 1 in 100,000, then the VSL is calculated as $1 million ($10*100,000). Thus the VSL is the value of reducing the probability that one anonymous person in a group dies. It is important to understand that the VSL is not the value of saving a certain person's life, but the value of saving an *anonymous* person's life. In policy applications, the benefit of a policy to reduce mortality is computed by multiplying the VSL by the reduction in deaths expected from the policy.

There are several methods used to estimate the VSL, and one of them is the hedonic wage method. The hedonic wage method uses labor market to analyze the individual's risk taking behavior and estimate the worker's marginal willingness to accept a marginal increase in risk. The basic idea behind the hedonic wage model is that in a competitive market, the worker who faces a higher level of disamenity on the job, such as fatal risk, must be compensated with higher wages than the worker who faces a lower level of disamenity on the job, or ceteris paribus.

Empirically, the hedonic wage model regresses the worker's occupational risk level as well as observed worker and job characteristics on wage. The estimated coefficient for the risk variable represents the additional wage workers require to accept an additional unit of risk, the wage-risk premium. This risk premium is then aggregated over the pool of workers at risk to estimate the value that workers collectively place on reducing the risk that one among them dies, which is equivalent to the VSL.

The literature of hedonic wage model is extensively reviewed in Fisher et al. (1989), Viscusi (1992) and Viscusi and Aldy (2003) as well as analyzed in Mrozek and Taylor (2002) and Kochi, Hubbell and Kramer (2006). Beginning with early work in the mid-1970s (Smith, 1974, 1976; Thaler & Rosen, 1976), there have been nearly 50 hedonic wage studies that estimate the VSL, with applications in many countries including the U.S., U.K., Canada, Australia, and some Asian countries (Kochi et al., 2006).[1]

Most hedonic wage studies use cross-section ordinary least squares (OLS) models to estimate the wage-risk premium. The VSL estimates from studies using data in U.S. workers have a quite wide range from \$0.1 million (Dilingham, 1985) to \$43.3 million (Olson, 1981) in 2005 dollars (Kochi et al. 2006). Mrozek and Taylor (2002) conduct a meta analysis of the VSL literature and conclude that the main source of variation in the past VSL estimates comes from differences in the quality of occupational risk data and

[1] This is the number of hedonic wage studies published until 2002. There are more recent studies published such as Black and Kniesner (2003), Viscusi (2004), and Kniesner et al. (2005).

samples of workers (blue collar or white collar workers or mix of both) and differences in the hedonic wage model specifications.

The U.S. Environmental Protection Agency (U.S.EPA) and some other government agencies use hedonic wage studies as a primary source of information for their VSL estimates. For example, according to the 2004 U.S.EPA White Paper regarding the use of VSL, the U.S.EPA used $4.8 million in 1990 dollars or $7.1 million in 2005 dollars as the VSL to assess their policy benefits (USEPA, 2004). This value was first estimated for the 1997 retrospective analysis of the Clean Air Act and is the average of 21 estimates from previous hedonic wage studies and 5 estimates from previous contingent valuation studies. All hedonic wage studies considered in 1997 retrospective analysis employed the cross-section hedonic wage models.

Recently, the U.S.EPA updated its VSL to $5.5 million in 1999 dollars (or $6.4 million in 2005 dollars) to evaluate the benefit of reducing mortality by Inter-State Air Quality Rule. This is the central estimate between the reported "best estimates" from two meta-analyses, Mrozek and Taylor (2002) and Viscusi and Aldy (2003), that analyze previous hedonic wage studies only (U.S.EPA 2004). The Office of Management and Budget (OMB) also uses the lower and upper bound of the "best estimate" from these same two meta-analyses of hedonic wage studies to evaluate the policy benefits for preventing the premature mortality when an individual agency does not monetized this benefit(OMB, 2003).[2]

[2] Informing Regulatory Decisions: 2004 Draft Report to Congress on the Costs and Benefits of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities retrieved March 25, 2007 from http://www.whitehouse.gov/omb/inforeg/draft_2004_cbreport.pdf .

The importance of the VSL for federal benefit cost analysis can be seen in *The Benefits and Costs of the Clean Air Act: 1970-1990* conducted by the U.S.EPA to report the periodic assessment of costs and benefits of the Clean Air Act to congress (U.S.EPA 1997). U.S.EPA (1997) reports that monetized benefit of the Clean Air Act during 1970-1990 is \$22 trillion with 5^{th} and 95^{th} percentile of \$5.6 and \$49.4 trillion, respectively (in 1990 dollars). The cost of the policy is estimated as \$0.5 trillion (in 1990 dollars). Over 80% of the benefits come from preventing premature mortality. The substantial uncertainty in the policy benefit estimation also comes from the large variance of the VSL estimate used where the mean VSL is 4.8 million with standard deviation of 3.2 million (in 1990 dollars).[3]

Given the strong link between the hedonic wage estimation and the policy evaluation, the quality of hedonic wage studies has been a major concern (U.S.EPA 2004). Although there have been significant improvements in hedonic wage literature, especially in the quality of fatal risk data, there are still many criticisms regarding the use of the hedonic wage model to evaluate certain government programs, such as environmental policy programs.

One of the major concerns regarding the hedonic wage model in policy analysis is the potential bias in the estimators due to omitted variables. A number of studies have indicated that the omission of unobserved individual heterogeneity, such as risk preferences or the worker's skill in protecting themselves in a dangerous work environment, are potentially important sources of bias in estimating the risk-wage premium (McConnell, 2006; Viscusi & Aldy, 2003).

---

[3] U.S.EPA apply Weibul distribution on the VSL estimates, so the range of VSL considered in the report is always positive.

There are several studies that addressed this problem. One group of studies uses an instrumental variable approach to obtain unbiased risk estimators in cross-section hedonic wage models (Arabsheibani & Marin, 2001; Garen, 1988; Gunderson & Hyatt, 2001; Siebert & Wei, 1994). The other group of studies uses panel models to control for unobserved time-invariant heterogeneity (Black, Galdo, & Lin, 2003; Brown, 1980; Kniesner, Viscusi, Woock, & Ziliak, 2005). The instrumental variable studies generally report unstable and unreliable risk estimators due to the weak instruments problem. The panel models show more robust results, but the estimators may be still significantly biased due to exacerbated measurement error bias and omitted time-variant heterogeneity.

The first objective of this dissertation is to identify the potential magnitude and direction of bias in previous cross-sectional hedonic wage studies. Unobserved time-invariant heterogeneity is controlled for by employing a first-difference model and a fixed-effect model on national panel data. The issues of exacerbated attenuation bias in the first-difference and fixed-effect model due to measurement error of the risk variable, as well as the bias due to the time-variant omitted variables and simultaneity between wage and risk variable are also identified by employing the instrumental variable approach on the panel models.

This dissertation is the first study to combine panel models and the instrumental variables approach to examine the comprehensive bias due to omitted variables in cross-sectional hedonic wage models. The results are used to evaluate past hedonic wage studies and potential biases that may have affected their wage-risk premium estimates.

Another concern regarding the use of hedonic wage estimators in policy analysis is the applicability of VSL estimates from labor market studies to policies with very different safety contexts, such as reducing the risk of death from pollution exposure (Hammitt, 2000; USEPA, 1997, 2004, 2005). The hedonic wage model estimates the tradeoff between an occupational risk and wage. Some psychology studies have found that individuals value the reduction of different types of risk differently (Slovic, Fischhoff, & Lichtenstein, 1980). Individuals may have a very different perception about occupational risk and other types of risk, such as environmental risk, and may place different monetary value on each type of risk reduction.

There are several stated preference studies examining the link between the risk characteristics and the individual's willingness to pay (Cookson, 2000; McDaniels, Kamlet, & Fischer, 1992; Subramanian & Cropper, 2000). However, there is little evidence regarding the nature of fatal risk/dollar tradeoff using revealed preference methods, such as hedonic wage models. Since the stated preference methods may contain the hypothetical bias, or bias resulting from difficulties in the communicating risk, it is important to verify whether or not individuals exhibit different willingness to pay to reduce different types of risks with revealed preference approach, which observes individual's actual behavior.

The second objective of this dissertation is to evaluate if workers have different willingness to pay to reduce different types of risks (e.g. violent assaults vs. traffic-related accidents) using a hedonic wage models. We use a specialized sample of workers combined with location and occupation-specific risk rates to determine if workers command different wage-risk premia for different types of occupational fatality risks.

More specifically, we examine whether fatal risks from violent assault risks (i.e., homicide) are compensated differently than non-violent risks.

A sample of occupational drivers is used, which includes truck, taxi, sales and bus drivers. These occupations are focused on for two reasons. First, these occupations face either high violent assault or high non-violent assault (or both) risks routinely as part of their job which may mitigate the measurement error due to the disparity between perceived risk and objective risk level. Second, these workers have very similar non-risk job characteristics which could help to mitigate bias due to unobserved job heterogeneity between occupations. Importantly, a location-specific fatal risk rate for two types of risks for each occupation is created. This allows risk to vary across Metropolitan Statistical Area (MSA) or state

This is one of a few studies that use hedonic wage studies to address the heterogeneity of risk and its effect on individual's willingness to pay to reduce a marginal amount of risk. Also, this is the first study that use location-occupation specific fatal risk rate. This study evaluate if the previous finding from the stated preference methods can be verified by the revealed preference method. Although the focus of this study is to estimate the different willingness to pay to reduce different types of *occupational* risks, the results of this study still can provide useful information whether or not the VSL used in the evaluation of policy which is different from occupational safety needs to be adjusted accordingly.

The remaining chapters are organized as follows. Chapter 2 reviews the theory of the hedonic wage model and important assumptions. The maintained hypothesis that we examine in this dissertation are also discussed. Chapter 3 presents data used in each estimation in later chapters. Two sets of worker data are used in this dissertation. In chapter 4, the Survey of Income and Program Participation, a national panel data, is used to conduct the panel data analysis. In chapter 5, a large-scale cross-section demographic data for occupational drivers is constructed from the Current Population Survey and is used to analyze the heterogeneity of risk preferences among workers.

We also use different risk data for each analysis. In chapter 4, we use an occupation-industry fatal risk matrix created by Scotton (2000). This risk data is varied by occupation and industry, which enables us to identify the risk change among workers who change jobs between or within industry and occupation. In chapter 5, we create an original occupation/location specific risk rate from the Census of Fatal Occupational Injuries by Bureau of Labor Statistics.

Chapter 4 presents the literature review and the analytical results of our panel data analysis of the hedonic wage model. We review past efforts to control unobserved heterogeneity in hedonic wage models. Then we describe the panel methods we employ, and present the results and conclusions. Chapter 5 presents the literature review and analytical results of our second objective of this study. We review previous stated preference studies to estimate different willingness to pay to reduce different types of risks, and discuss their potential problems. We also review the previous hedonic wage study which we improve upon. We present the estimating hedonic wage models and the robustness of our estimation. Lastly, we present our conclusions. Chapter 6 discusses the findings in the main analyses, and its contribution to the VSL and hedonic wage literature. We also discuss the policy implications of our results and future research.

Chapter II

Theory of Compensating Wage Differentials

This chapter reviews the theory of compensating wage differentials (hedonic wage theory) as related to measuring the value of reducing mortality risks. The theory of hedonic wages is largely based on the theory of hedonic pricing developed by Rosen (1974). Rosen analyzes the price determining process of attributes of goods in the implicit market**.** Jones-Lee (1974) also analyzes the workers risk-wage taking behavior using the state dependent expected utility theory. In general, both theories from Rosen and Jones-Lee are combined to describe the underlying theory of hedonic wage.[4]

In a hedonic wage equilibrium, there is assumed to exist a hedonic wage schedule, v, which relates all relevant job characteristics to wages. The job characteristics of interest here is risk, r. Thus the hedonic wage functions is written, $v(r,z)$, where z subsumes all job characteristics other than risk. This hedonic wage schedule is an envelope function arising from *all* worker's utility maximization and *all* firm's profit maximization process. For each individual worker or individual firm, the market hedonic wage schedule is given. Workers maximize their utilities and firms maximize their profits subject to a given hedonic wage schedule.

In the following sections, we base on Jones-Lee (1974), Rosen (1974) and McConnell (2006) for the description of hedonic wage theory. As mentioned earlier, Jones-Lee developed the basic theory of worker's expected utility maximization between the states of survival and death, and Rosen developed the general theory of hedonic

---

[4] See also reviews of hedonic wage theory in McConnell (2006) and Scotton (2000).

pricing models. McConnell (2006) provides the overview of hedonic wage models incorporating more recent findings from hedonic wage literature.

After describing the basic hedonic wage theory, we discuss important underlying assumptions in the hedonic wage model. We discuss the assumption of homogeneity among workers or firms and the homogeneity of fatal risk. There are two aspects of the hedonic wage literature as related to estimating the VSL that we wish to extend in this dissertation.

Preferences over Risks and Wages: the Worker

In this section, we review the workers risk-wage taking behavior using the state dependent expected utility theory developed by Jones-Lee (1974). Given that workers face the uncertain outcomes of death and survival at work, we assume that the individual worker maximizes expected utility between utility in the death state and utility in the survival state. First, we define state-dependent utility over these two states. Let $U_S(W)$ be utility in the survival state, which depends on the level of wealth (wage). We assume that utility is increasing at a decreasing rate as wealth increases, which denotes

$$\frac{\partial U_L(W)}{\partial W} = U_L' > 0 \tag{1}$$

and

$$\frac{\partial U_L(W)^2}{\partial^2 W} = U_L'' < 0\,. \tag{2}$$

We also define $U_D(W)$ as the utility in the death state. We assume that utility in the death state is also a function of wealth (wage), since individual may obtain a certain utility by bequesting the wealth to heirs. If this is not the case, then the utility in death state is zero. Thus, we define

$$\frac{\partial U_D(W)}{\partial W} = U_D' \geq 0 \tag{3}$$

and

$$\frac{\partial U_D(W)^2}{\partial^2 W} = U_D'' \leq 0\,. \tag{4}$$

We also assume that for a same level of wealth, workers obtain higher utility in the state of survival than in the state of death, thus

$$U_S > U_D. \tag{5}$$

Workers choose a job with a probability of death, r, to maximize their expected utility. Suppose workers initial wealth is W (>0), and the initial level of fatal risk is $\bar{r}$ ($0< \bar{r} <1$). Consequently their expected utility is:

$$EU_0 = (1-\bar{r})U_S(W) + \bar{r}U_D(W) \tag{6}$$

Now, suppose workers fatal risk level is increased to r ($0< \bar{r} < r <1$). To sustain the original level of utility, workers need to be compensated by additional wages. Let the compensation level that leaves workers in the same expected utility level be V. We call V the willingness to accept (WTA) compensation for an increment change of risk. We assume that V is a function of the risk preference parameter, α. By definition, the following equality holds:

$$(1-\bar{r})U_S(W) + \bar{r}U_D(W) = (1-r)U_S(W+V(\alpha)) + rU_D(W+V(\alpha)) \tag{7}$$

The left hand side is a fixed level of expected utility, say EU_0. By rearranging (7), we obtain the function for V at the fixed level of expected utility, EU_0, which is

$$V = V(r, W, \alpha, EU_0)$$

Substituting back this V function to (7), and by total differentiating with respect to r, the marginal willingness to accept compensation for a marginal change in risk, $\partial V/\partial r$, is given by:

$$\frac{\partial V}{\partial r} = \frac{U_S(W+V) - U_D(W+V)}{(1-r)U_S' + rU_D'} \tag{8}$$

where $U_s' = \partial U_s(W+V)/\partial V$ and $U_D' = \partial U_D(W+V)/\partial V$. From (1), (3), (5), the equation (8) is positive. The second derivative of (8) is also positive, indicating that the workers WTA curves are positively sloped and convex.

Figure 1 shows the mapping of two individuals' WTA curves. To simplify the argument, we assume that individuals are equally productive. The wage level is determined by the worker's productivity as discussed detail in the next section. The variation of identically productive workers wage may arise when workers choose different risk levels.

Figure 1 maps two WTA functions, where V_1 is the WTA locus for worker 1 and V_2 is the WTA locus for worker 2. Along each WTA locus, the utility level is constant. The workers WTA is a function of risk level, r, worker's risk preference parameter, α, and wealth level, W. Since the WTA function is convex, the movement towards north-west direction increases their utility level. In figure 1, worker 1 has a higher utility level in V_1' than V_1. In addition, worker 1 is more risk averse than worker 2, thus the V_1 curve is steeper than V_2 curve. The steeper V curve indicates that the worker 1 requires more wage compensation to accept a marginal increase in risk than worker 2 when they face a same risk level. The workers' different marginal rates of substitution between risk and compensation force workers to find different optimal levels of risk-wage compensation when they face a same market constraint. Workers maximize their utility level where V curves and hedonic wage schedule are tangent to each other.

Firm's Production and Risk

A firm's profit function can be complicated depending on which cost factors we include (McConnell, 2006). Here we assume a simple profit function that is:

$$\pi = \varphi z(L) - wL - c(r;\mu)L \text{ ,}^{5} \tag{9}$$

where φ is the price of output, *z(L)* is the production function as defined over labor inputs, w is the wage rate, c is the cost of providing a certain level of safety per worker and μ is the efficiency parameter of providing a certain safety level. For simplicity, we assume there is no factor other than labor to produce output. Firms maximize their profits subject to the hedonic wage schedule (thus $w=w(p;z)$) by choosing the number of workers hired and the level of safety provided.

The profit maximizing conditions respect to L and r are:

$$\frac{\partial z}{\partial L} = \frac{w+c}{\varphi} \tag{10}$$

$$\frac{\partial w}{\partial r} = -\frac{\partial c}{\partial r}. \tag{11}$$

Equation (10) shows that in order to maximize profits, the firm should choose labor inputs such that the marginal productivity of labor (left hand side) should equal to the wage rate plus the cost to provide safety per worker (right hand side) assuming the price of output equal to one. Equation (11) shows that the firm's marginal implicit price

[5] This is a slightly modified profit function presented in McConnell (2006), which assumes that firms do not pay death benefit and compensation for the event of injury, and available instant replacement of an identical worker when the worker die. See McConnell for the discussion with more comprehensive profit function.

of risk represented by the hedonic wage schedule (left hand side) should be equal to the marginal cost-saving from a marginal increase of risk level (right hand side).

By rearranging firm's profit function (9), we obtain the firm's wage offer curve (OC);

$$OC = w(r, L, \mu, \pi) = [\varphi z(L) - c(r;\mu)L - \pi]/L. \qquad (12)$$

This offer curve shows the firm's trade-off between providing safety and wage compensation for a given profit level (thus this is an isoprofit curve). The slope of OC curve is given by $\partial OC/\partial r = -\partial c/\partial r$. Assuming the cost of providing additional unit of safety is decreasing at a decreasing rate as risk increases,[6] OC is a positively sloped concave function from below. At the maximum profit level given hedonic wage function, the conditions (10) and (11) hold.

Figure 1 illustrates the OC for two types of firms. To simplify the argument we assume all firms offer identical job characteristics except fatal risk level. The only difference comes from the different efficiency of providing the safety at work. Firm 1 has an offer curve 1 (OC_1) and Firm 2 has an offer curve 2 (OC_2). Firms' profits are higher on offer curves that are in the south-eastern direction. For example, firm 1's profit level is higher in $OC_1{}'$ than OC_1 since at every risk level, firm 1 pays less wage for workers in OC1' than OC1. Firm 1 has a flatter offer curve than firm 2, indicating firm 1 has a higher efficiency of providing an additional safety than firm 2. Firms maximize their profit where their offer curves are tangent to the hedonic wage schedule. The different shapes of offer curves force firms to find different optimal levels of providing

[6]The cost incurred by reducing a marginal risk at high risk levels are lower than the cost incurred by reducing a marginal risk at a low risk level.

safety and compensation given the market equilibrium wage/risk constraint. Firm 1 provides higher safety level than firm 2, given the hedonic wage schedule.

Figure 1 also shows how the hedonic wage schedule emerges when there are heterogeneous workers and heterogeneous firms. As described more detail in McConnell (2006), the hedonic wage schedule is generated by the joint distribution of workers' risk preference (α) and firms' efficiency to provide safety (μ) holding other job and worker characteristics constant. Worker 1, who is more risk averse than worker 2, will be hired by firm 1 for whom providing safety is relatively inexpensive at a risk of r_1 an wage of $wage_1$. Worker 2, who is less risk averse than worker 1, will be hired by firm 2 for whom providing safety is relatively costly. The hedonic wage schedule is an envelope function which traces out the each worker/firm labor contracts, where the workers' marginal willing to pay for a marginal decrease in risk level is equal to the firms' marginal cost of providing a marginal decrease in the risk level.

As noted in Rosen (1974) and further explained by McConnell (2006), the hedonic schedule is solely determined by the distribution of firms' offer curves in the long run. In the long run, due to the entries into the market, firms' profits are fixed at zero. Provided that in a large economy there is a continuum of workers' preferences over risks and wages, the distribution of firms' profit functions (or envelope profit schedule) reflects the hedonic wage schedule. This is important because as seen later, it relaxes the information requirements to recover the marginal willingness to accept by workers to take more risky jobs. This envelope profit schedule will shift in two cases (McConnell, 2006). One case is when technology innovation alters firms' cost functions to supply safety. The other case is when the distribution of risk preference of workers changes. Redistribution of risk preference affects the distribution of isoprofit functions through changes in the wage level and the profit levels of firms.

Underlying Assumptions

The major underlying assumptions of the hedonic wage model are that the labor market is competitive, highly mobile, and the workers and firms have perfect information regarding the occupational risk level (McConnell, 2006). However, the perfect information assumption may be relaxed in the long run. It may not be reasonable to assume that the workers have perfect information about the occupational risk level of a particular job they are about to take. Workers would have a certain level of perceived fatal risk level of that job, and use this perceived risk level to determine their optimal wage/risk compensation. This perceived fatal risk levels may or may not be the same as actual risk levels.

Researchers do not observe the workers' perceived risk levels, and thus use the objective level of risk on the job, which is estimated from occupational fatal statistics. If the labor market is in a long-run equilibrium, the use of objective risk levels would result in unbiased estimates of wage/risk tradeoffs if firms perceive risks accurately. In the long run, there is an entry in the market and the firms profit level is fixed at zero. This constraint results in the hedonic wage schedule being the envelope of firms' isoprofit functions, regardless of the perceived risk level by workers. Thus researchers observe risk-wage pairs that are on the hedonic wage function and are able to recover an unbiased hedonic wage function.

In the short run, however, there is no entry into the market and firms can enjoy positive profits. In this case, the hedonic wage schedule is influenced by both workers and firms. When all workers underestimate the fatal risk level, the estimated hedonic wage functions is also underestimated. This is illustrated in figure 2. Suppose there is "actual" hedonic wage function which is determined by the "actual" risk level. Worker 1 has an V_1 and he perceives the risk level of the job at $risk_{perceived}$ and demands wage compensation at $wage_0$. However, suppose that the actual risk level for this job is $risk_{actual}$ where $risk_{perceived} < risk_{actual}$. Assume that a firm knows the actual risk level, and is willing to pay up to $wage_1$ as compensation. At the end of wage negotiation, the wage level will likely be determined between $wage_1$ and $wage_0$ levels. Since we only observe the actual risk level, we match the wage level (likely) less than $wage_1$ and $risk_{actual}$ as opposed to $wage_1$ and $risk_{actual}$ to estimate the hedonic wage function. This causes an underestimation of the actual hedonic wage schedule. Similarly, if all workers overestimate the risk level, we are likely overestimating the hedonic wage function. If there are mix of workers who underestimate or overestimate risk level, the bias in the hedonic wage estimation depends on the distribution of worker's mis-perception of risk levels.

Another important assumption of the hedonic wage model is homogeneity among workers in their productivity and job characteristics. The hedonic wage schedule is determined by the characteristics of job and the characteristics of workers (McConnell, 2006). As stated in the basic theory of hedonic wage section, we assume identical workers in terms of their productivity. The only difference among workers is the difference in their risk preference. The same is true for firms. We assume firms are identical except their efficiency in providing safety.

When there are heterogeneous groups of workers or firms, then each group would have different hedonic wage schedules. For example, assume that college graduates and high school graduates have an identical distribution over risk preference parameter. Generally college graduates are more productive than high school graduates because of their higher human capital. Even though the marginal cost to provide safety is same for both groups at every risk level, firms would pay college graduates who have higher marginal productivity more than high school graduates. Thus college graduates and high school graduates have different hedonic wage schedules as illustrated in figure 3. This is a particularly important point in estimating hedonic wage model empirically. Failing to take into account differing productivities of workers (i.e., the slope of the hedonic wage schedule between different groups of workers) may cause biased estimates of the HW schedule and could even cause negative slope estimates in the hedonic wage function (McConnell, 2006).

In terms of job characteristics, homogeneity in terms of firms benefit package would be particularly important to control for in estimating a HW function (McConnell, 2006). If one group of firms offer full employer provided health insurance plan and another group of firms offer no health insurance plan, then their wage schedules should be different. The former group of firms can attract workers with fewer wages than latter group of firms.

We also generally assume that worker's risk preference does not change for heterogeneous risks. Even for fatal risk, there can be heterogeneity among risks depending on the circumstances of death. As discussed in detail in chapter 5, previous studies indicate that individuals may have different risk preferences towards different types of risks. The different risk preferences seem to depend on the different qualitative characteristics of risk such as level of controllability or dread involved. If workers consider different types of risks as separate job characteristics, or marginal cost of reducing different types of risk are differed, then there exists separate hedonic wage schedules for different types of risks.

Extensions Examined in This Dissertation

In implementing the theory and estimating hedonic wage functions, a number of key assumptions are invoked. Chapter 4 examines the implication of failing to take into account all heterogeneity among workers and firms in the empirical analysis. As reviewed by McConnell (2006), it is common to control for the heterogeneous levels of productivity (actual or perceived) among workers resulting from market segmentation based on factors such as race, gender, and education level in the empirical analysis. However, it is less common to take into account the heterogeneity among jobs. In addition, there may be more heterogeneous levels of productivity among workers originated from workers characteristics which are unobservable to researcher.

Chapter 4 examines the effect of the unobserved heterogeneity among workers on estimating hedonic wage model. As discussed in chapter 4, we employ panel models that control for all time invariant worker characteristics. We also explore the effect of unobserved time-variant worker characteristics and the effect of workers' potential mis-perceptions about the risk levels they face on the job. While we focus on unobserved worker heterogeneity, we also attempt to improve on earlier studies, with regards to firm heterogeneity. We include several important job characteristics in our estimation model, such as employer provided health insurance availability and firm size. The firm size may represent the overall amenity of firms. Unfortunately, a lack of data on firms precludes a more sophisticated analysis.

The second assumption underlying most hedonic wage estimates is that all risks may be aggregated into a single event; death, even though the circumstance of each death can be very different. The distribution of workers' risk preferences and firms' efficiency to supply safety may be significantly different among different circumstances of fatal accidents as discussed earlier. If this is the case, there may be separate HW functions for each type of risk, and estimating a HW function with an aggregated measure of risk may significantly bias the estimation. This hypothesis is formerly examined in Scotton and Taylor (2006). As presented by Scotton and Taylor (2006), when there are heterogeneous risks, the workers expected utility (6) becomes:

$$EU = r_N U_S(W) + \sum_{i=1}^{N-1} r_i U_{Di}(W) \tag{13}$$

where r_N is the probability of survival (r_N >0), r_i is the probability of death by the fatal event i (r_i >0), $r_N + \sum_{i=1}^{N-1} r_i = 1$, U_s is the utility level in the survival state, and U_{Di} is the utility in the death state resulted in the fatal event i. The fatal event, i, may be car accident, exposure to toxic materials, homicide and so on. Since U_{Di} is a separate utility function for each event of death, the compensation required to accept a marginal increase of each type of fatal risk may differ significantly. In chapter 5, we improve the study design upon Scotton and Taylor (2006) and estimate separate hedonic wage functions for different types of risks to examine the second assumption of homogeneity among fatal risk.

The data considerations for each of the above extensions, as well as a review of the relevant literature and the testable hypotheses we develop, are presented in detail in chapters 3, 4 and 5. In the next chapter, the data available to explore our extensions to previous work estimating hedonic wage functions are presented.

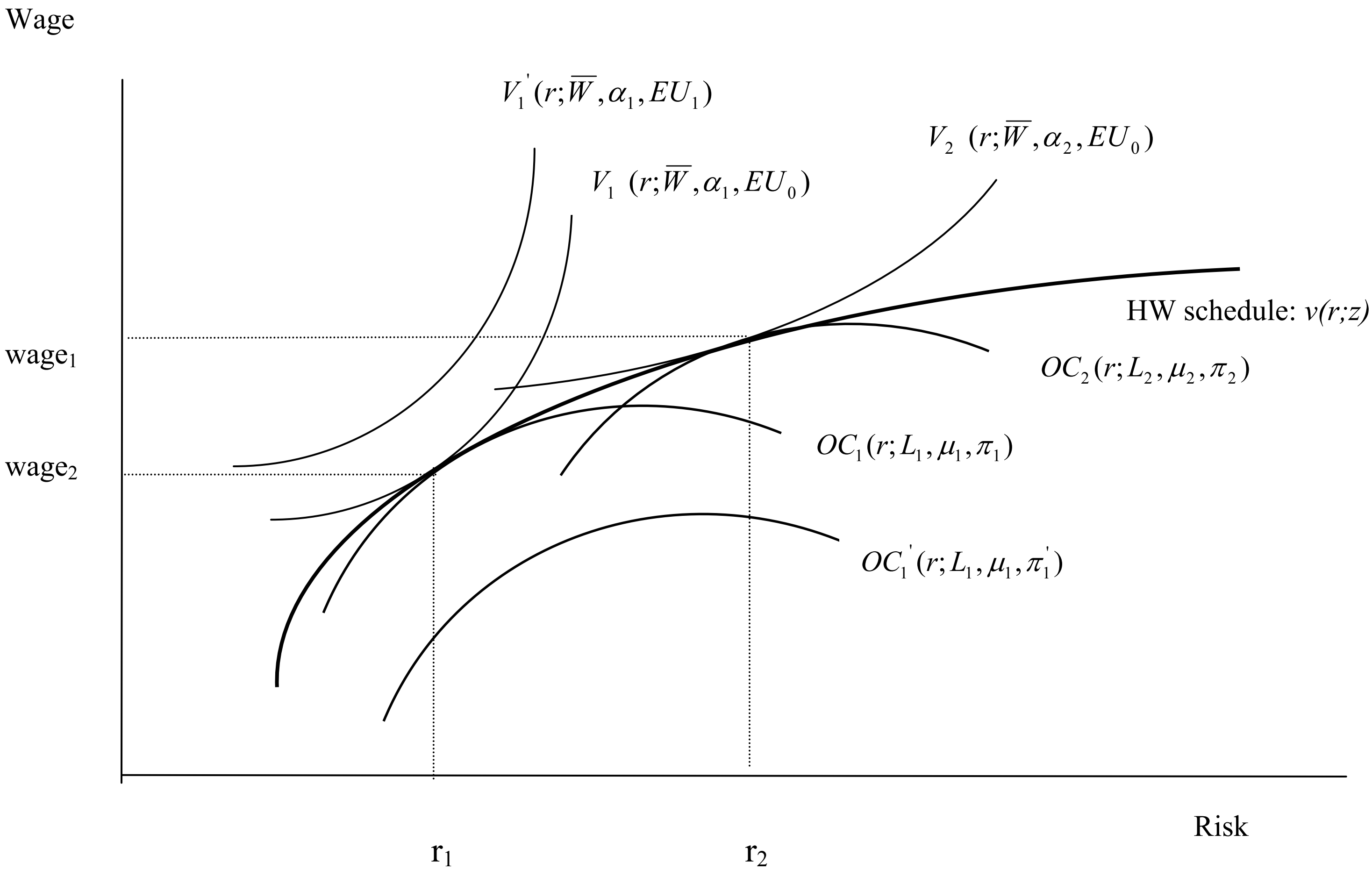

Figure 1. HW equilibrium schedule for heterogeneous workers and heterogeneous firms.

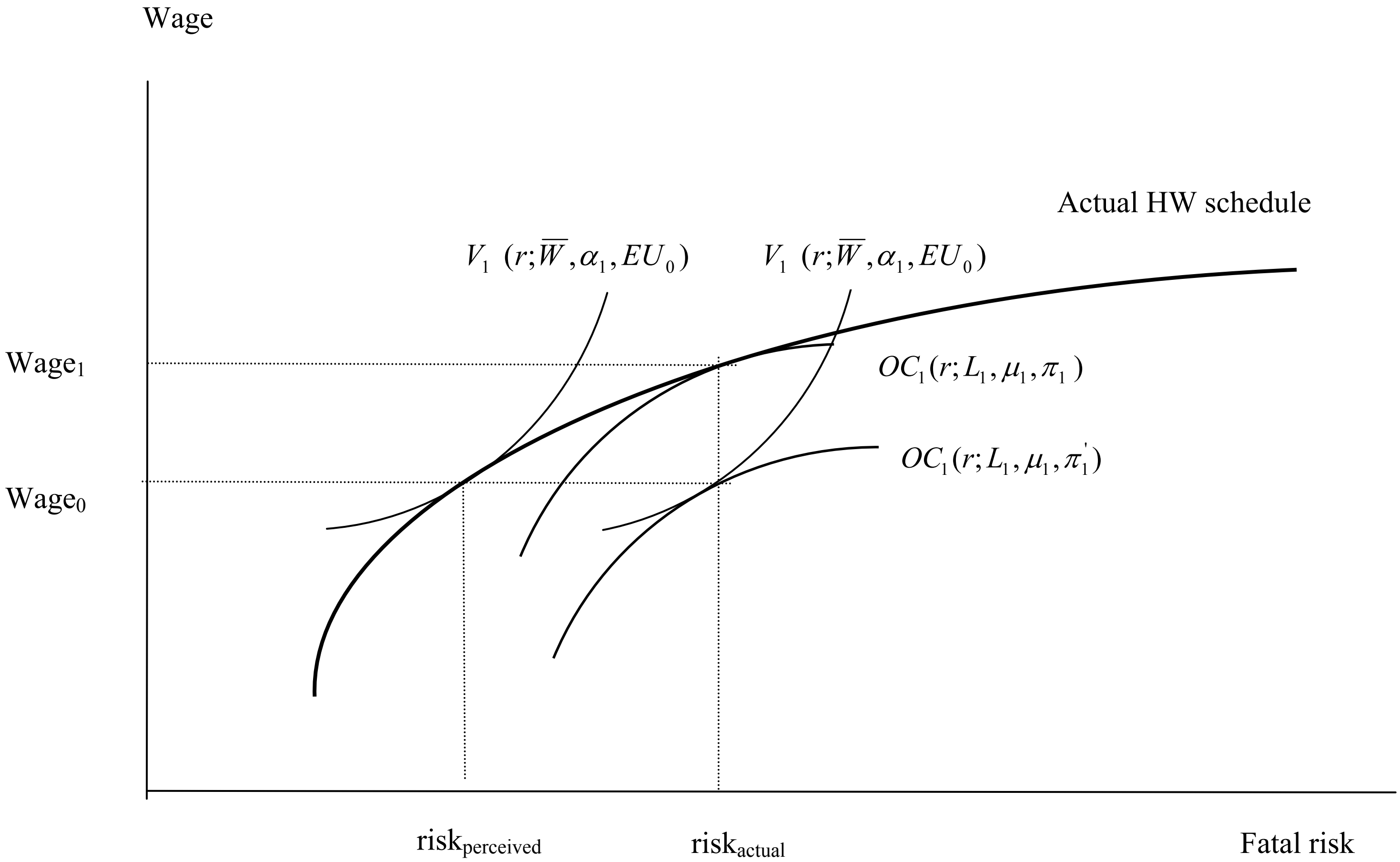

Figure 2. HW equilibrium schedule with workers who underestimate risk levels.

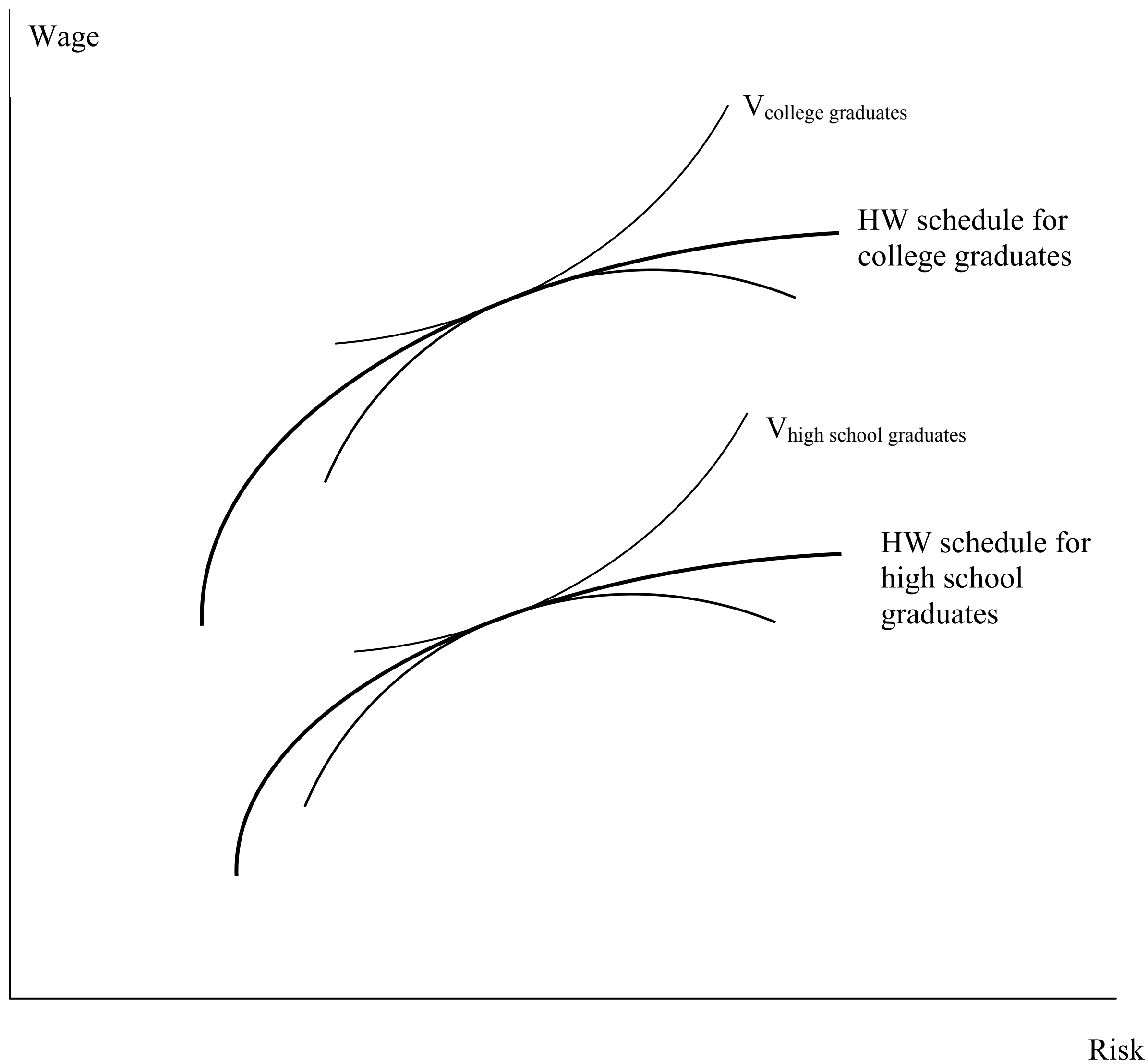

Figure 3. HW equilibrium schedules for workers with different productivity.

Chapter III

Data

This section describes the data used in this dissertation. The key data needed to estimate a hedonic wage (HW) model are occupational fatal risk and labor force data. Due to their different focus, chapter 4 and chapter 5 will rely on different data sets to estimate hedonic wage models.

Chapter 4 focuses on controlling for worker heterogeneity in a hedonic wage model through panel data methods. For this purpose, the demographic information must come from longitudinal data. The Survey of Income and Program Participation (SIPP) is used to construct this longitudinal data. It is also critical to identify the changes in risk level when workers change their jobs. An industry based risk level, which has been most commonly used in hedonic wage literature, is not a good candidate for this purpose since it does not capture the changes in risk if workers change occupations within the same industry group. The risk data should at least vary by occupation and industry. Occupational fatal risk data created by Scotton (2000), which varies by reasonably disaggregated occupational and industry codes is used.

Chapter 5 focuses on identifying potential heterogeneity in risk-wage premia for different types of risk. To accomplish this, a sample of occupational drivers is used. These workers are chosen to control heterogeneous aspects of job requirements as well as mitigate measurement error associated with the risk variable. In a cross section study such as that employed in chapter 5, it is difficult to control for unobserved non-risk aspects of jobs, such as working conditions and job requirements, even though these could be important factors to determine the wage, as well as the risk level. Using a sample of occupations which requires similar working conditions and job skills may enable us to mitigate the potential endogeneity problem associated with omitted variables in HW models. In addition, there is likely a disparity between the objective risk measure and the subjective risk measure, which can be another source of endogeneity problems in the HW model. Benjamin (2001) observed that the individuals who are facing the high objective risk often understand their objective risk level correctly. Driving is one of the riskiest jobs in the United States. For example, in 2003, 861 sales workers and truck drivers died at work. This is the highest level of death among all occupation groups.[7] The fatality rate of these drivers is 2.67 per 10,000 workers, which is the fourth highest risk rate among all occupations.[8]

The demographic information for this sample comes from the basic monthly Current Population Survey collected by the BLS. We divide risks into two types for this analysis: violent assault risks and all other risks, which mostly relate to traffic accidents. These risks for occupational drivers largely vary by geographic area (Knestaut, 1997). Thus the risk data varies by occupation and geographic area. This occupational fatal risk data is created from the Census of Fatal Occupational Injuries (CFOI) 1992-2002,[9] and the Occupational Employment Statistics 1998-2003, both collected by the BLS. In addition, various geographic specific wage determinants, such as the unemployment rate,

[7] Census of Fatal Occupational Injuries 2003 data retrieved March 25, 2007 from http://www.bls.gov/iif/oshwc/cfoi/cfch0002.pdf
[8] Ranking of high risk occupation is (from the highest to lowest): logging, aircraft pilots and farmers and ranchers.
[9] This research was conducted with restricted access to Bureau of Labor Statistics (BLS) data. The views expressed here do not necessarily reflect the views of the BLS.

local industry composition, population, local total vehicle miles traveled (VMT) are collected.

The remainder of this chapter is organized as follows. First, the demographic data and risk data used for chapter 4 are discussed separately. Then, demographic data are merged with the risk data and risks for the sample of workers used in chapter 5 are presented. Data for chapter 5 are then discussed, following the same format as the discussion for the data used in chapter 4.

Data for Panel Study (Chapter 4)

Labor force data: the Survey of Income and Program Participation (SIPP)

Data for individual hourly wage, job and socio-economic characteristics come from the Survey of Income and Program Participation (SIPP).[10] The SIPP is national panel data administered by the U.S. Census Bureau. The SIPP contains rich information about individual income, labor force status, and general demographic characteristics of U.S. population. People are interviewed by phone or in person every four months. Each four months reference period is called a *wave.*[11] Only one observation from each wave is used.

The 1996 SIPP data is used in this analysis.[12] The 1996 SIPP panel lasts for four years and contains twelve waves. The total sample sizes in the beginning of the 1996 panel are approximately 115,700 individuals.[13]

---

[10] Detailed data description and data download is available at http://www.bls.census.gov/sipp/ (retrieved March 25, 2007)

[11] Some questionnaires ask people to record information for every month since the last interview. In this case, only the fourth observation is used for the analysis.

[12] The SIPP is available for 1988, 1990, 1991, 1992, 1993, 1996 and 2001 panels. Samples in each panel are independent from each other. Since the risk data is available only from 1992 as discussed later, the 1990 and 1991 panel are removed from the analysis. The 2001 panel is also removed because the risk data is available only until 1997 and there is no overlap in the time period between the 2001 SIPP panel and the risk data. The 1993 SIPP panel does not include many eligible workers who has two or more observations,

The main advantage of using the SIPP data as compared to other labor force data such as the CPS and the PSID is its structure as panel data and the richness of the information it contains. The SIPP is a medium length time series data with large sample size. The CPS is a relatively short length time series data with limited wage observations. In the CPS, for each individual, there are only two observations obtainable for the current wage level, while there are nine to twelve observations on current wage levels in the SIPP. The PSID is the long length time series data (since 1968), but with relatively small sample size (about 6,000-8,000 households).

As compared to the CPS and the PSID, the SIPP also provides more information on employer characteristics, as well as current earned and unearned income. The SIPP collects important employer characteristics such as the availability of employer-provided health insurance and the size of firms. According to the SIPP User Guide,[14] the SIPP collects 70 cash and in-kind source income data for the current year, [15] while the CPS collects only 35 and the PSID collects only 25 cash and in-kind source income for the *prior calendar year*.

The major wage determinants commonly used in the hedonic wage literature are available in the SIPP. These include age, educational attainment, gender, race, marital

---

and thus is also removed from analysis. The data from the 1992 panel is not used in this analysis due to a concern about the data quality.

[13] The sample size of SIPP is increased in 1996 panel dramatically as a part of SIPP reform. This reform is called 1996 redesign. Main aspects of reform include: lager sample size, longer panel, introduction of computer-assisted interviewing system, which automatically check the consistency of reported data during the interview, and oversampling the low-income household (http://www.sipp.census.gov/sipp/evol.html, retrieved March 25, 2007).

[14] Available via http://www.sipp.census.gov/sipp/usrguide/sipp2001.pdf (retrieved March 25, 2007).

[15] Cash and in-kind source income include: wage, earnings from various financial and real estate investments, and payment relating to the workers compensation, such as temporal sickness benefits, pensions and government welfare programs.

status, number of children under 18 in the household,[16] union status, and the occupation and industry group of the firm for which workers work. The worker's current wage is available on either an hourly or monthly basis. The hourly wage data is preferred since there are more missing observations in the monthly wage data. Also collected are residential location (urban vs. rural), region, availability of employer provided health insurance, size of firms, and whether or not the person works over-time as potentially important wage determinants.

Table 1 shows the definition of variables extracted from the 1996 SIPP panel and the summary statistics for the sample of workers. The sample is full time workers[17] who hold only one job at a time. The workers who are earning less than minimum wage,[18] or whose age are less than 18 or more than 65 are omitted from the analysis. There are total of 166,362 observations for 34,846 workers. The minimum, average and maximum number of observations per worker is 1, 4.8 and 12, respectively. The average hourly wage is $13.27, which is lower than the average hourly earning in the U.S. labor market that is $16 in 2005.[19] The average age of workers is 38 years old, and 41% of the sample graduated from high school, 33% of the sample has attended college, and 9 % of the sample has a bachelors or higher degree. The current national trend is that approximately 30% of the workforce having graduated from high school, 30% have attended a college but have no degree, and 30% hold a bachelor's degree or more.[20] Compared to the current

---

[16] Inclusion of the number of children in a HW equation is rather rare, but is seen in Lenoie et al. (1995), Herzog and Schlottman (1990), Siebert and Wei (1994) and Sandy and Elliot (1996).
[17] Full time workers are defined as workers who work more than 35 hours per week.
[18] The minimum wage level for the service workers ($2.13 per hour) is used as a cut off wage level.
[19] October 2005 Employment Situation Summary retrieved March 25, 2007 from http://www.bls.gov/news.release/empsit.nr0.htm. All monetary values are adjusted to 2005 dollars using the consumer price index.
[20] The educational attainment level of labor force over time retrieved March 25, 2007 from http://www.bls.gov/cps/labor2005/chart2-1.pdf.

national trend of educational attainment in the U.S. labor force, our sample under-represents the labor force with bachelor's degree or more, and over-represents the labor force with less than high school diploma.[21]

About 45% of the sample is female and 56% of the sample is married. The majority of workers are white, 13% of workers are Hispanic and another 13% of workers are African American, which correctly reflects the recent racial composition of the labor force in the US.[22] About 19% of workers are union members or covered by union benefits, which is slightly higher than the current average union membership rate in the U.S. of 12.5%.[23] The average number of children for workers in the sample whose age is less than 18 is 0.79, which is less than the national average of 0.9.[24] More than 60% of workers receive part or full health insurance through their employers. Seventy-eight percent of workers lived in urban areas. There is a slightly larger proportion of workers in the South, and a smaller proportion of workers from the Northeastern region. Fifty-six percent of workers work in the firms with more than 100 employees in all location, and 29% of workers work in the firms with less than 25 employees in the worker's location.

The share of workers who work in each major occupation and industry group is following.[25] Only five percent of workers have a farming related occupation. Twenty-eight percent of workers have a technical occupation, 26% of workers have a labor

---

[21] This difference may come from the oversampling of low-income household in SIPP.
[22] Selected labor force characteristics of Hispanics or Latinos retrieved March 25, 2007 from http://www.bls.gov/cps/labor2005/chart4-2.pdf. According to this, the Hispanic makes up 13% of labor force in 2005.
[23] Union members summary: Union members in 2006 retrieved March 25, 2007 from http://www.bls.gov/news.release/union2.nr0.htm.
[24] This is the average number of children under 18 per household. Average number of children per family and per family with children, 2000 Census retrieved March 25, 2007 from http://www.census.gov/population/socdemo/hh-fam/tabST-F1-2000.pdf.
[25] Major industry and occupation classification follows Scotton (2000), and is reproduced in appendix B and C, respectively.

occupation, 14% of workers have a service occupation, 16% of workers have a craftsman occupation, and 11% of worker has a professional occupation. There are few workers engaging in the agricultural, construction, transportation/ communications/utility and public industry. These workers makes-up only 15% of total sample. In contrast, a high proportion of workers engage in the trade, service and manufacturing industries. Thirty one percent of workers engage in the service industry, 19% of workers engage in the wholesale or retail trade industry, and 25% of workers engage in the manufacturing industry.

In the 1996 SIPP sample, there are total 16,001 workers who change their occupation or industry at the 3-digit level classification group between waves. This is approximately 9% of total observations in the sample. Among these job changers, 5,849 workers change the occupation within the same 3-digit level industry group, 1,587 workers change the industry but stay in the same 3-digit level occupation group, and 8,565 workers change both the occupation and the industry group. Although there may be significant number of workers who change their jobs within a same occupation and industry, the exact number who does so is difficult to identify. The SIPP provides information on whether or not a worker changes jobs between consecutive waves. However, many individuals in our sample do not have observations in consecutive waves because an observation(s) from a previous or following wave is dropped due to missing data, or because they are ineligible to be included.[26] Thus without a balanced panel from

---

[26] For instance, say a worker was employed full time in one job in wave 1. In wave 2, the worker takes a second part-time job and keeps this for two waves. This worker would not be included in wave 2 or 3 because sample is limited to those working one fulltime jobs. If in wave 4, we observe the worker in new job (as compared t wave 1) and working one fulltime job again, they would re-enter our sample.

consecutive waves, the exact number of workers who change their jobs within the same occupation and industry is uncertain.

Table 2 shows the summary of job changing behavior between major occupation groups among the 1996 SIPP sample. The first row shows the major occupation groups workers belong to at t where t=2, 3,…12. The first column shows the major occupation groups workers belong to at t-1. For example, the second column-third row cell represents the number of workers who change job from a craftsman occupation to a professional occupation. About 5-10% of workers in each occupation group change jobs to different major occupation groups. Although workers move between all major occupation groups, there are tendencies for workers in certain occupations to change jobs to certain other occupations. For example, relatively high proportion of workers who are in a professional occupation switch to a technical occupation, and the workers who are in a technical occupation tend to switch to a professional, labor or service occupation. In general, there is a balanced in-flow and out-flow in each occupational group during the panel.

Table 3 shows the summary of job changing behavior between major industry groups among the 1996 SIPP sample. There are 3-12% of workers in each industry group who switch their jobs to a different industry group. The workers who are in the public industry has the lowest rate to switch jobs to a different industry (3.05%), and the workers in the construction industry has the highest rate to switch jobs to a different industry (12.14%). There are also tendencies of workers job changing pattern between major industry groups. The workers in a construction industry tend to switch jobs to an agriculture, service or manufacturing industry. Workers in a wholesales trade, service or manufacturing industry tend to switch jobs among these industries. The workers in a transportation/communication/utility industry tend to switch to a service or wholesales trade industry. The workers in a public industry tend to switch to a service industry.

Changes in key demographic characteristics of workers (regardless of whether they change jobs) are as follows. There are 1,503 times that someone changes marital status from single to married and 805 times that someone changes from married to single.[27] These make-up about 0.9% and 0.4% of total observations, respectively. There are 2,358 observations that become new parents during the panel, and there are 3,341 observations whose underage kids become older than 18, or deceased during the panel. There are 3,660 observations that have a union membership at t-1, and lose that membership at t, after they change jobs. There are 3,553 observations who newly acquire union membership due to a job change.

Occupational fatal risk data.

Scotton (2000) creates 506 risk rates based on a 22 occupation × 23 industry fatal risk-rate matrix. To avoid measurement error due to yearly fluctuations of death incidences, Scotton computes a six year average risk rate between 1992 and 1997. The risk rate in each occupation-industry cell is calculated by the following formula:

$$p_{oi} = \frac{D_{oi}}{W_{oi}} \tag{14}$$

where p_{oi} is the fatal risk rate in occupation o and industry i, D_{oi} is the annual average number of death incidents in occupation o in industry j, and W_{oi} is the annual average total number of workers in occupation o in industry i.

[27] An individual may changes marital status more than one time. Thus the number of changes may be greater than the number of individuals who change marital status.

The numerator in equation 14, D_{oi}, is obtained from the CFOI files for the period 1992-1997, which contains more than 37,000 deaths. CFOI for this period uses 3-digit occupation code from the Census Occupation Classification System 1990, and 4-digit Standard Industrial Classification (SIC) 1987 code to classify the occupation and industry, respectively (Scotton 2000). There are 473 occupation and 1,183 industry categories included in 1992-1997 CFOI. The list of all variables available in the CFOI files is reproduced in Appendix A. The CFOI contains various characteristics of deceased workers, such as gender, age, race, location of accident, size of firm, event of accident, occupation, industry, and time of accident.

Scotton regroups occupation and industry codes in the CFOI into an original matrix of 22 occupation and 23 industry codes. She obtains the annual average deaths in each occupation industry pair between 1992 and 1997. In the equation 14, W_{oi} is obtained from taking annual averages of employment levels in the industry and occupation pairs from the Industry-Occupation Employment Matrix (OEM) 1991-1996 administered by the BLS. Scotton's 22 occupation and 23 industry classification is reproduced in Appendix D and E. Appendix D shows the occupation group mapping. The first column shows Scotton's 22 occupation group, the second column is the title of the occupational group and the third column is the census' 3-digit occupation categories which are included in each occupation group created by Scotton. Appendix E shows the industry group mapping. The first and second columns show the Scotton's 23 industry group title and codes, respectively. The third and fourth columns show the 2-digit SIC title and codes included in Scotton's grouping, respectively. The last column shows the corresponding industry code in the SIPP for each of the 23 industry groups.

The SIPP uses the same 3-digit occupation code as the CFOI, so the occupation code in the SIPP is directly converted to 22 occupation groups created by Scotton. The SIPP uses the 3-digit industry code following the 1990 Census classification as described in the SIPP data dictionary.[28] The data dictionary also shows the corresponding 4-digit SIC code for each 1990 Census classification system industry categories. Scotton presents a corresponding 2-digit SIC code for her 23 industry group. Thus for this analysis, the SIPP industry codes are matched to the 2-digit SIC codes, and then converted to the 23 industry groups used by Scotton.

First, we briefly discuss the average annual risk rate among the 506 occupation-industry groups. The average number of deaths is nine in every 100,000 workers or 0.9×10^{-4}. The highest risk bearing group is the construction tradesmen in the personal transportation service industry, where the risk is 35.5×10^{-4}. The second highest risk bearing group is the agricultural workers in the lumber/wood/stone/glass product industry where the risk is 12.4×10^{-4}. High risk bearing *industries* include the personal transportation industry and the mining industry. High risk bearing *occupations* include construction tradesmen and truck drivers. Examples of a low risk bearing *industry* is the social/legal/education service industry, while and example of a low risk bearing *occupation* is the financial record keepers. Fifty-one occupation-industry groups had no deaths during the 6 year period.

After merging the risk data with the SIPP workforce data, we can examine the distribution of risk in our sample. The mean fatal rate in the SIPP sample is 5.5×10^{-5} with standard deviation of 9.5×10^{-5} and the median risk rate is 2.0×10^{-5}. This is comparable to the mean risk rate of related studies which use the occupation within industry risk rates.

[28] Data dictionary is available via http://www.bls.census.gov/sipp/diction.html (retrieved March 25, 2007).

Scotton (2000) reports a mean risk rate for her sample from CPS (n=4,891) is 5.0×10^{-5} with standard deviation of 9.0×10^{-5}. Kniesner et al.(2005) report a mean risk rate for their sample from the PSID (n=7,937) is approximately 5.8×10^{-5} with a standard deviation of 8.5×10^{-5}.

About 0.6% of total observations, 1,156 observations in the SIPP sample, face zero risk. There are only two observations in the SIPP sample that face the highest risk level which is 35.5×10^{-4}. Both workers are construction tradesmen in the bus service and urban transit industry. The second highest risk level is 12.4×10^{-4}, and there are 75 observations who face this level of risk.

There are a total 13,733 observations in the SIPP sample where the risk rate changes between waves due to worker's job changes. This comprises approximately 8% of total observations. The risk change ranges from -12.3×10^{-4} to 12.4×10^{-4}. The workers who experience the largest negative risk change, -12.3×10^{-4}, are those who change the job from a timber cutter in the millwork industry to a production coordinator in the same industry. The workers who experience the largest positive risk change, 12.4×10^{-4}, is the worker who change the job from a stock/inventory clerk in the health service industry to a timber cutter in the logging industry.

The mean risk change is 0.02×10^{-5}. Table 4 and 5 shows the mean risk change between major occupation and industry groups, respectively. The structure of tables is same as table 2 and 3. The risk change is the risk level at t minus the risk level at t-1. Thus the positive (negative) number indicates the increased (reduced) risk level due to a job change. Diagonal entries are not zero since there can be some risk changes within an aggregated level occupation/industry group if workers change jobs at more disaggregated occupation/industry groups. As shown in table 4, the mean risk change within the same major occupation group is small and it ranges from -0.4×10^{-7} to 0.4×10^{-6}. However, for some occupation groups, the range of risk changes is large even for workers who stay in a same occupation group. For example, craft, technical, and labor occupations show risk changes ranging between -5.0×10^{-4} and 5.0×10^{-4}. Considering that the mean risk level of workers in the sample is 0.5×10^{-4}, this risk change is significant. The variation of risk changes within the same major occupation group comes from the changes of industry group and the changes of occupation group at a disaggregated level. On the other hand, professional and service workers experience relatively small risk change when they stay in the same occupation groups, ranging between -1.0×10^{-4} and 1.0×10^{-4}, and -1.5×10^{-4} and 1.5×10^{-4}, respectively.

On average, for technical workers, switching jobs to any other major occupation group increases their risk level. For labor, farming, and craft workers, switching jobs to any major occupation groups except farming or labor reduces their occupational risk level. For the professional workers, switching jobs to technical occupation reduces their risk level. For service workers, changing the job to the professional or technical occupation reduces their risk level.

The job changing pattern between major industry groups is as follows. On average, the workers in the service industry increase their risk level when they change the job to any other industry. For workers in the construction, agriculture or transportation/communication/utility industry group, changing the job to other industries other than these three major group industries reduces their risk level. The workers moving into the service industry from other industries reduce their risk level, and the workers moving into the construction industry increase their risk level.

Table 6 shows the summary changes in the risk level of workers who do not change demographic categories and table 7 shows the summary changes in the risk level of workers who change demographic characteristics. For example, suppose worker 1's marital status is single from wave 1 to 10 and is married in wave 11 and 12. Changes in risk levels between wave 1 and wave 2, wave 2 and wave 3, …, and wave 9 and wave 10 (total 9 observations) for worker 1 are included under "single worker" category in table 6. The change in risk level between wave 11 and wave 12 (one observation) is included under the "married worker" category in table 6. In other words, the change in risk levels between wave 10 and wave 11 where worker 1 changes his/her marital status is not included in table 6 but included in table 7 under "single to married" category. Demographic categories examined here include, gender, marital status, kids status, and union status.

As shown in table 6, for male workers, the mean risk change level is 3.9×10^{-7} while that of the female workers is 0.2×10^{-8}. Due to the large standard deviation, these mean values are not significantly different from zero. In fact, none of demographic category has a mean risk change value which is different from zero. Almost 90% of female or male workers have zero risk change and thus the median risk change level is zero. This high proportion of zero risk changers is also true for all other demographic categories. Once zero risk changers are removed, we have higher mean risk change values of 3.7×10^{-6} and 0.2×10^{-7} for male and female workers, respectively. The variance of risk change for the male workers is much larger than that for the female workers as well. This does not change after we remove zero risk changers. This may be due to the limited availability of high risk jobs for female workers while male workers have more mobility between safe and risky jobs.

For workers who remain single and workers who remain married between waves, the mean risk changes are 0.3×10^{-6} and 0.1×10^{-6} respectively (including zero risk changers). There is not much difference between these two groups of workers in terms of standard deviation, maximum and minimum value of risk change for both with and without zero risk changers. For workers whose kids' status does not change between waves, the mean risk changes are 0.1×10^{-6} and 0.9×10^{-7} for those with underage kids and those without underage kids, respectively (including zero risk changers). In general, workers without kids have more variation in risk change than workers with kids.[29] This may indicate that the having underage kids forces workers to make a conservative decision in terms of occupational risk when they change a career.[30] For the workers who remain as a union member, and workers who remain as a non-union member between waves, the mean risk changes are 0.2×10^{-6} and 0.1×10^{-6}, respectively. The workers who remain as a non-union member show wider variation of risk changes than union-members. Although table 6 shows the smallest risk change for unionized worker is -10.74×10^{-4}, there is only one observation for this level of risk change. The second smallest risk change for unionized workers is -5.64×10^{-4}, thus the more reasonable range

[29] Although the table 6 indicates the similar min and max risk change of workers with underage kids compared those of workers without underage kids, there are only one observation at upper and lower bound level of risk changes. Removal of these observations reduce the range of risk change to -5.64×10^{-4} to 5.50×10^{-4}.

[30] Of course there may be other demographic characteristics which explain the difference of the risk change range between two groups. For example, workers with underage kids are younger and more likely being married than workers without underage kids.

of risk change for these workers is -5.64×10^{-4} to 4.96×10^{-4} which is much narrower than the risk change for non-union workers (-12.33×10^{-4} to 12.41×10^{-4}).

This section describes the risk level change associated with the changes in workers demographic characteristics shown in table 7. In general, the range of risk change for workers who change demographic characteristics is narrower than that of workers who do not change demographic status.[31] [32] The difference in the range of risk change between these two groups of workers may be due to the different sample size. For example, there are only 2,308 worker-wave observations who change marital status while there are 131,515 worker-wave observations that do not change marital status. The small sample size in table 7 may fail to capture the workers movement between extreme risk changes. Risk changes seem to be randomly distributed on observable demographic characteristics changes as indicated the mean changes not being significantly different from zero. This is consistent with the mean risk changes for workers whose demographic status does not change. Interestingly, however, workers tend to change risk level when they change a demographic status more often than when staying in a same demographic category. The proportions of zero risk changers among workers who stay in a same demographic status are between 4 and 11% depending on the category, while that among workers who change a demographic status are between 17 and 28%.

[31] Although the maximum risk change for workers who change marital status from married to single is 11.37×10^{-4}, there is only one observation that experiences this level of risk change. The second highest risk change is 4.31×10^{-4} for these workers.

[32] Although the table 3.7 indicates that the min and max risk change of workers who change from non-union to union status is -11.58×10^{-4} to 10.83×10^{-4} there are only one observation at upper and lower bound level of risk changes. Removal of these observations reduces the range of risk change to -5.36×10^{-4} to 5.58×10^{-4}.

Data for Chapter 5

This section is organized as follows. First, the labor force data used in chapter 5, the Current Population Survey (CPS), is discussed. Also discussed are potential wage determinants for drivers that are not available from the CPS, such as regional characteristics including the local annual unemployment rate, and the per capita sales volume of wholesale, retail, transportation, and entertainment industries. Next section discusses the occupational fatal data followed by the employment data and the summary of risk rates for each occupational driver. Also the non-fatal injury risk data and the summary of non-fatal injury risk rates for each occupational driver is discussed. Lastly the summary statistics for fatal and non-fatal events for the sample workers are discussed.

Labor force data: Current Population Survey (CPS).

The Current Population Survey (CPS) is a national survey administered by the Census Bureau and the BLS. Each survey includes about 50,000 households. It is a monthly survey and each respondent participates for 16 month. The samples are interviewed each month for the first four months, and then take a break for eight months, and then come back and are interviewed each month for a final four months. The sample used in the analysis is group of respondents who are either in the fourth or the eighth "Month in Survey," in other word, in the fourth or the 16th month of their participation period. Survey participants in the fourth or 16th month are referred to as the "outgoing rotation group." The outgoing rotation group is a preferred sample because of the availability of *current* wage data. We collect the sample of workers every other year so that no person appears in an outgoing rotation group more than once.

The sample of drivers is obtained from the CPS administered in 1994, 1996, 1998, 2000 and 2002. The sample is limited to a non self-employed, single job holding, full time occupational drivers. Workers earning less than minimum wage[33] and workers in Hawaii and Alaska are not included. Self employed workers are omitted for several reasons: 1) our focus is to estimate hedonic wage equations resulting in the wage-risk negotiation, however this would not be the case for self-employed workers, and 2) as described later, the risk measure does not reflect the risk of self employed workers.

There are total 19,371 occupational drivers in the CPS sample including truck, sales, bus and taxi drivers. The worker's occupation in the CPS is coded according to the Census Occupation Classification, which follows the Standard Occupation Classification 1980 definition. In Census Occupation Classification, the following codes are assigned to each occupational driver; 804 for truck drivers, 806 for sales drivers, 808 for bus drivers and 809 for taxi drivers. The detailed definition of each occupational driver is presented in Appendix F. Truck drivers operate tractors, heavy or light trucks. They transport, deliver or/and pick up goods and merchandise. Sales drivers operate trucks or other vehicles to deliver, sell, or collect goods over establish routes. Bus drivers transport passengers inter-city and intra-city by bus. Taxi drivers operate automobiles or limousines to transport passengers.

The wage, individual and job characteristics other than occupational fatal risk (X) are also obtained from the CPS. The variables in X_i includes age, educational attainment, race, U.S. citizenship, gender, usual work hours, union status, marital status, location of

[33] Minimum hourly wage for taxi driver is \$2.13 per hour. Minimum wage per hour for other drivers is \$4.25 in 1994, \$4.75 in 1996, and \$5.15 in 1998, 2000 and 2002. Per hour minimum wage is multiplied by 35 hours to obtain weekly minimum wage.

household, occupation dummy, regional dummy and MSA dummy variables. Table 8 shows the definitions, data source and summary statistics of each variable.

The average weekly wage of occupational drivers is $683 which is higher than the average U.S. worker's weekly wage which is about $530 and the SIPP sample used in chapter 4.[34] In chapter 5, all monetary values are adjusted to 2004 dollars using the consumer price index. Average age of the sample is 40 years old which is slightly older than the SIPP sample. About half of the sample graduated from high school, 23% of the sample received a college level education, and 4% of the sample graduated from a four year college. Compared to the SIPP sample, the educational attainment level is slightly lower among the CPS driver sample, and compared to the national average, the educational attainment level is even lower.

Ten percent of the sample is of Hispanic origin, and 12% is African American. Most of the sample has U.S. citizenship (93%) and only 6% of the sample is female, reflecting the male-dominance of the occupations upon which we focus. Fifty-percent of the sample works overtime and 23% of the sample holds a union membership, both of which are noticeably larger than that of our general sample from the SIPP used in chapter 4. In the SIPP sample, only 19% worked overtime and 19% held a union membership, which is still higher than the national average (12.5%). Sixty four percent of the sample is married, which is again higher than the proportion of married sample in the SIPP (56%). In summary, occupational drivers tend to be older, less educated, highly male dominated, engaged in more overtime work, and more heavily unionized than the national average and our general SIPP sample from chapter 4.

[34] According to October 2005 Employment Situation Summary published by BLS. Monetary value of average weekly wage for the U.S. worker is adjusted to 2004 currency level by consumer price index.

Local area unemployment statistics.

The local unemployment rate is a potentially important factor for determining the wage level, as well as the willingness of workers to accept workplace risk. The Metropolitan Statistical Area (MSA) unemployment rate in 1994, 1996, 1998, 2000 and 2002 is obtained from the Local Area Unemployment Statistics (LAUS) administered by the BLS. [35] The average annual unemployment rate during the sample period (1994-2002) is between 4% and 6%. However, the unemployment rate varies significantly across MSAs. For example, some MSAs such as the McAllen-Edinburg-Mission MSA in Texas recorded quite high unemployment rates (over 10%) between 1994 and 2002, while others such as Madison, Wisconsin experienced quite low unemployment rates (2-3%) during the sample period. In general, the West census region experienced relatively high unemployment rate (average 6%) compared to other regions (4.5% to 5.5%) during the sample period.

Economic activity in each MSA.

Different levels of economic activity in each MSA could affect the wage level of workers through different levels of demand for occupational drivers. The volume of sales in wholesales, retail, transportation, arts-entertainment, and accommodation and food service industries in each MSA would likely affect the demand for the truck, sales, bus and taxi drivers. These data are obtained from the 1997 Economic Census administered by the U.S. Census Bureau and are only available for 1997.[36] The population of each

[35] The data is available via the Local Area Unemployment Statistics (LAUS) retrieved March 25, 2007 from http://www.bls.gov/lau/home.htm.
[36] The data is available via 1997 Economic Census retrieved March 25, 2007 from http://www.census.gov/epcd/www/econ97.html.

MSA is also obtained from the 1997 Economic Census to calculate the per-capita sales volume and employment level in the above industries.

The vehicle miles traveled (VMT).

The number of vehicle miles traveled (VMT) in each MSA may indicate different traffic levels and workload for occupational drivers among MSAs. This difference may affect the wage level of occupational drivers, and the fatal or non-fatal risk levels the drivers face at work. The VMT is obtained from Bluestone (Forthcoming). Bluestone collects the county level VMT in 1996 from the U.S. Environmental Protection Agency (EPA)[37] and aggregates it to the MSA level. The VMT is computed on a per-capita basis.

Occupational fatal incidences.

This section describes the source of occupational fatal incidence data. The next section presents the local employment data used to calculate the local fatal risk rates, followed by a discussion of the estimated local fatal risk rates.

The number of occupational fatal incidences is obtained from the Census of Fatal Occupational Injuries (CFOI) file collected by the BLS for the period 1992 through 2002. This data differs from that in chapter 4 by level of detail and is thus not publicly available. The non-public use CFOI file contains information on the location of injury at the county level in addition to the information available in the public CFOI (see appendix A). For this research, the location of injury at the county level is aggregated into an MSA level count of deaths using the 1999 MSA definition by the Office of Management

[37] U.S.EPA (1998) National Air Pollutant Emission Trends Update, 1970-1997," EPA-454/E-98-007, U.S. Environmental Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, December 1998

and Budget.[38] This is the definition used in the 1999-2002 Occupational Employment Statistics (OES), which provide the source of our employment level in each MSA.[39] Later, the count of deaths in each MSA for each occupational driver is divided by the employment level of each occupational driver in each MSA, so the MSA definition should be matched between denominator and numerator. The location of injury at the state level is directly coded from the CFOI.[40]

Two types of fatal events are also obtained directly from the CFOI; homicide and non-homicide deaths. For ease of exposition, we refer to homicides as violent assault deaths and non-homicides as non-violent deaths. Violent assaults include assaults and violent acts by persons such as hitting and shooting, and do not include self-inflicted injuries and assaults by animals. Non-violent events include deaths from all sources other than homicide and self-inflicted injuries.

There are a total of 10,475 non-self employed occupational drivers deaths during the period 1992 to 2002 or 952 deaths per year on average.[41] Among them, 8,872 are deaths of truck drivers, 468 are sales drivers, 198 are bus drivers and 937 are taxi drivers. The number of deaths is re-counted for each occupational driver in each MSA/state for each death event. The summary of the fatal incidence for each type of driver for each event is shown in table 9. The transportation related injuries are the main cause of death for truck, sales, and bus drivers and accounts for 70 to 80% of total deaths. Violent assaults are the main cause of death for taxi drivers, accounting for 70% of total deaths.

[38] MSA definition available from Metropolitan Areas and Components 1999 with FIPS code retrieved March 25, 2007 from http://www.census.gov/population/estimates/metro-city/99mfips.txt.
[39] We also use 1998 and 2003 OES which use slightly different MSA definitions. However, definition change between 1998 OES, 2003 OES and 1999-2002 OES is marginal and should not have any effect on using 1999-2002 MSA definition for entire period.
[40] Both county and state are coded based on FIPS. In CFOI, state of New York is coded as 68 instead of 36 as of FIPS.
[41] This number includes the deaths due to self-inflicted injury and assaults by animals.

Other deaths include being struck by objects, caught in equipment, compressed or pinched by rolling, sliding or shifting objects, caught in or crushed in collapsing materials, falls, bodily reactions and exertion, contact with electric current, exposure to temperature extremes or to caustic, noxious or allergenic substances, ingestion of substances, fire, or explosion.

The majority of violent assault deaths occur inside MSAs. Of those who die of a violent assault, 84 percent of truck drivers, 87% of sales drivers and 91 % of taxi drivers died within an MSA. The number of death within an MSA for bus drivers and all drivers are suppressed for reasons of confidentiality. A high level of deaths by homicide within MSAs could be due to the higher crime rates in urban areas as compared to non-urban areas. Non-violent deaths including transportation related deaths and other types of deaths often occur outside MSAs. Of those who die of a transportation related event, 47 percent of truck drivers, 36% of sales drivers, 35% of bus drivers, and 21% of taxi drivers died in transportation related events outside MSAs. Also, of those who die of other types of death, 36% of truck drivers and 31% of sales drivers died outside the MSAs. In the analysis, the regression results using different combinations of MSA-level and state-level violent and non-violent risk rates will be presented.

Table 10 reports the MSAs where the top five *numbers of deaths* occurs for truck, sales and taxi drivers summarized by the event of death. The second and the fourth columns show the actual number of deaths in each MSA with a corresponding risk rate, which discussed later. Los Angeles and Chicago are ranked in the top five for occupational deaths in each driving occupation regardless the death event. The other cities in the top five ranking are large MSAs, such as Dallas, Washington, DC, Miami, Atlanta, Houston, Philadelphia, Detroit, and New York. Not surprisingly, New York City has an outstanding number of violent assault deaths of taxi drivers. One-hundred, eighty-six taxi drivers (17 taxi drivers annually) died due to the violent assault over the last 11 years. Also table 11 shows the list of states where the top five *number of deaths* occurs for the truck, sales and taxi drivers. For truck drivers, California, Texas and Florida have the highest deaths for both violent and non-violent events. For sales drivers, Pennsylvania, Texas and Georgia record the highest deaths among all states for both events. For taxi drivers, New York, California, Florida and Illinois are the states with the highest level of deaths regardless of the event.

Area specific employment level: Occupational Employment Statistics (OES).

The fatal risk rate is calculated by equation 14. The annual average number of deaths in each occupation and in each MSA/state is divided by the annual average employment level in each occupation and in each MSA/state. The number of workers in each driving occupation in each MSA/state is collected from the Occupational Employment Statistics (OES) administered in 1998-2003 at the MSA level and 1998-2004 OES at the state level by the BLS. Unfortunately, employment data at the MSA and state level (for detailed occupations) is not available prior to 1998. In addition, the OES does not include self-employed workers, therefore self-employed worker deaths are removed from the risk estimation.

There can be significant bias of our risk estimation due to the use of data from a different time period in the numerator (period in 1992-2002) and the denominator (period in 1998–2003). If the actual employment level between 1992 and 1997 is significantly lower (higher) than the employment level between 1998 and 2003, the estimated risk rate is underestimated (overestimated) as the annual average risk rate of the period in 1992-2002.

Table 12 is the assessment of the potential bias due to the limited employment data period. According to the BLS's Current Employment Statistics (CES), which report a longer period of national employment data, there is the following employment trend for the driving occupations between 1992-1997 and 1998-2003. In the CES, there is no distinction between the truck driver and the sales drivers, so both occupations are combined into the "truck" category in table 12. The taxi and bus drivers are reported to have about a 3 to 4% increase in employment levels, while the truck drivers show a rather significant increase in the employment level between two periods (12%). This indicates that using the 1998-2003 average employment level as a denominator would likely underestimate the average fatal risk level of truck drivers between 1992 and 2002 periods.

A solution to this problem is to estimate risk rates using only the 1998-2002 death record to match the data period of the denominator and the numerator. A potential problem of this solution is a reduction of number of deaths due to the shorter period of data. The reduction in the number of deaths will generate more geographic areas with zero risk, which can lead to a less variation in the risk variables. This is a particular concern for the violent assault fatal risk.

Table 13 shows the summary of deaths by occupation in the period of 1998-2002. If the frequencies of deaths are constant over time, the number of deaths in the period of 1998-2002 should account for about 45% of total number of deaths between 1992 and 2002. The parentheses in table 13 shows the proportion of deaths occurred in the period

1998-2002 in the total deaths occurred between 1992 and 2002. For the violent assault case, truck drivers and taxi drivers have somewhat fewer deaths than expected during 1998-2002 while the number of transportation related deaths is increased for all drivers, particularly for bus drivers compare to the period of 1992-1997.

To further examine the potential impact of the available data on our hedonic results, hedonic wage models are estimated separately using two risk rates. First the risk rates created with the 1992-2002 death record are matched to our 1994-2002 worker data, and then the risk rates created with the 1998-2002 death record are matched to 1996-2002 worker data. For the model using the risk rate based on the 1998-2002 death record, the 1994 CPS is dropped from the analysis because its labor data period is far before the risk data period.

Fatal risk rate.

The summary of risk rates by occupation for each event is described in this section. Table 14 shows the mean and the standard deviation of the risk rates created from the 1992-2002 death record for each event of death and for each occupational driver. The risk rates are estimated both at MSA and state levels. Standard deviations are presented in parentheses. At the MSA level, the average violent assault fatal risk is 0.97 in 10,000 drivers or 0.97×10^{-4}, and the average non-violent fatal risk is 1.74×10^{-4}. Compared to the average risk level of all occupations, which is 0.9×10^{-4}, driving is a somewhat higher risk occupation. Taxi drivers report the highest violent fatal risk rate of 3.52×10^{-4}. The average violent fatal risk for the sales, bus and truck drivers are 0.24×10^{-4}, 0.13×10^{-4} and 0.06×10^{-4}, respectively. On the other hand, truck drivers have the highest non-violent fatal risk rate of 2.62×10^{-4}. Taxi drivers have the second highest fatal risk rate, 2.17×10^{-4}, followed by that of sales drivers of 1.02×10^{-4}, and that of bus drivers of 0.95×10^{-4}.

Compared to the MSA level risk, the mean value of state level violent assault risk is higher for all occupational drivers, and the mean value of state level non-violent risk is slightly higher or similar. For example, the mean state-level violent assault risk is 46% more than the mean MSA level risk, and the mean state level non-violent fatal risk is 14% more than the mean MSA level non-violent fatal risk. The largest difference in the mean violent risk can be seen for the taxi drivers where the state level risk is 51% more than the mean MSA level risk. The largest difference in the mean non-violent risk is for the bus drivers where the state level risk is 52% more than the mean MSA level risk which indicates that the deaths outside MSA are proportionally more than the number of workers outside MSA.

Table 15 shows the mean and the standard deviation of the MSA level fatal risk rate by occupation and event after dropping MSAs with less than 100 employees. Since some MSAs have very high risk rate due to their small employment level, we examine if the average risk level change by excluding MSAs with less than 100 employees. There is no change in the risk rate for the truck drivers. There are slight decreases in risk rates for other drivers, except the violent assault risk rate for the taxi drivers. The violent assault risk rate for the taxi drivers actually increased. This is because dropping MSAs with less than 100 employees removes many MSAs with zero risk as well as MSAs with a very high violent assault risk.

Table 16 and 17 repeat the information in table 14 and 15, but risks are created using only the 1998-2002 CFOI data. As indicated in table 16, the violent fatal risk rate created based on the 1998-2002 CFOI data is generally smaller than that created based on the 1992-2002 CFOI data except for the bus drivers. This is true for both the MSA level risk rate and the state level risk rate. For example, with the MSA level risk rate, the mean 1998-2002 CFOI based violent risk rate is 0.75×10^{-4} for all drivers, which is 22% less compared to the 1992-2002 CFOI based risk rate. The non-violent fatal risk created from the 1998-2002 CFOI is larger than that created from the 1992-2002 CFOI for the MSA level risk for any types of driver. It is slightly larger for truck and sales drivers and smaller for bus and taxi drivers when we compare the state level risk.

Table 17 is the mean and standard deviation of the fatal risk rate by occupation and the event of death based on the 1998-2002 CFOI data after dropping MSAs with less than 100 employees. There is no change in the risk rate for truck drivers. The risk level is decreased for other drivers except for the violent risk for the sales drivers.

The decrease in violent fatal risk rates based on the 1998-2002 CFOI data (except bus drivers) as compared to the risk rates based on the 1992-2002 CFOI data is as expected. The annual frequency of violent fatal incidence during the period of 1998-2002 is less than that of 1992-2002 (except bus drivers), but we use a same annual employment level to estimate both risk rates. For the same reason, the increase in non-violent fatal risk rates based on the 1998-2002 CFOI data as compared to the risk rates based on the 1992-2002 CFOI data is as expected.

Non-fatal injury incidence: Injuries, Illnesses, and Fatalities (IIF) program.

This section describes the number of non-fatal occupational injuries and its risk rates among occupational drivers. The non-fatal occupational injury risk is called *injury risk* in this study. The count of non-fatal injuries comes from the Injuries, Illnesses, and

Fatalities (IIF) program, administered by the BLS. The IIF provides information about the number of incidences of occupational injuries for each state in the United States. The number of injuries can be broken down into the following categories: state where the incidence occurs, employer profile (private, state government, or local government), year of incidence, and injured worker's characteristics such as industry and occupation affiliation and event of injuries. The aggregated number of injury cases for state-occupation-event pairs is readily available in the IIF CD-Rom (available from the BLS upon request). However, for each state-occupation-event pair, the information about industry affiliation of injured worker is not available. Thus, we include school bus drivers, which are removed from the fatal risk calculation, to estimate a injury risk for bus drivers.

The number of injuries is defined as the number of workers who experience days away from work due to an occupational injury. Occupational injuries which result in death are omitted from these statistics. The annual non-fatal injury data are available between 1992 and 2002 for most states. The more detailed data availability by state is summarized in table 18. Colorado, Washington DC, Idaho, New Hampshire, North Dakota and Ohio do not participate in the IIF program and so there is no injury data available for these states. The nonfatal injuries are divided into two events following the fatal risk creation: those caused by violent interactions such as worker conflicts or assaults and those caused by non-violent acts.

For each occupation, the number of nonfatal injuries for each event is divided by the number of total employment in each state to create non-fatal injury risk rates. The number of total employment is obtained from the OES. The injury risk is only available at the state level. For the same reason of the fatal risk case, the injury risk rate is calculated separately using 1992-2002 data and 1998-2002 data.

Table 19 shows the mean injury risk rates for each driving occupation across state for total risk, violent interactions and transportation related event created from the 1992-2002 injury data. Truck driving has the highest non-fatal injury rate for both total injury and transportation related injuries. Five in 100 truck drivers experience some kind of injury every year, and seven in 1,000 truck drivers experience some type of traffic-related injury. Bus drivers show the highest violent-related injury rate among all driving occupations. Four in 1,000 bus drivers suffer a violent-related injury every year. Taxi drivers experience a relatively low <u>non-fatal</u> violent-related injury (seven in 10,000). A relatively low <u>non-fatal</u> violent-related injury rate combined with a high violent-related <u>fatality</u> risk indicates that when a taxi driver is assaulted, it tends to be fatal.

For the truck, sales and bus drivers, the proportion of transportation related non-fatal injuries to total injuries is between 6 and 21%. This is a dramatic difference compared to the proportion of transportation-related fatal injuries of between 70 and 80%. This disparity of event component between fatal and nonfatal injuries may explain why there is generally no effect of adding injury risk on the fatal risk coefficient in the previous HW studies (Kochi et al., 2006; Mrozek & Taylor, 2002).

Table 20 shows the non-fatal injury rates by occupation based on the IFF 1998-2002. When the injury risk is created from the 1998-2002 injury data to match the period of employment data, the injury risk rate is generally slightly reduced as compared to the injury risk rate created based on the IFF 1992-2002.

Fatal and non-fatal risks in the sample.

This section describe the fatal and non-fatal injury risk level our sample of workers face. First I describes the risk rates created from the 1992-2002 CFOI, and then describe the risk rates created from the 1998-2002 CFOI. Table 21 shows the summary statistics of the MSA level risk rates based on the 1992-2002 CFOI in the sample. There are 12,892 total drivers who live in MSAs in the sample. Among them, 40% of drivers face zero violent fatal risk and 8% of drivers face zero non-violent fatal risk. The mean violent assault fatal risk rate is 0.83×10^{-4} and the mean non-violent fatal risk rate is 1.92×10^{-4}. Compared to the mean fatal risk level in the SIPP sample in chapter 4, the fatal risk level of occupational drivers is quite high.

The proportion of the sample living in MSAs which face zero violent fatal risk is 39% for truck drivers, 60% for sales drivers, 79% for bus drivers, and 14% for taxi drivers. On the other hand, there is generally a much smaller proportion of the sample which faces zero non-violent fatal risk (except for taxi drivers). Only 1% of truck drivers, 38% of sales drivers, 39% of bus drivers and 20% of taxi drivers face zero non-violent fatal risk.

The mean violent fatal risk rate for truck, sales, bus and taxi drivers is 0.08×10^{-4}, 0.30×10^{-4}, 0.18×10^{-4}, and 10.67×10^{-4} respectively. The mean non-violent fatal risk rate for truck, sales, bus and taxi drivers is 2.07×10^{-4}, 0.70×10^{-4}, 0.87×10^{-4}, and 2.38×10^{-4} respectively. Table 22 summarizes the state level risk rates in the sample that is constructed from the CFOI 1992-2002. In general, the state level violent assault risk is lower than MSA level violent risk, while the state level non-violent risk is higher than the MSA level non-violent risk for the sample. The number of workers with zero risk at the state level is much smaller than at the MSA level. With the state level risk rates, there are 18% of drivers who face zero violent fatal risk and only 1% of drivers who face zero non-violent fatal risk. There is no state in which truck drivers face zero non-violent fatal risk in the sample.

Table 23 and 24 shows a summary of the MSA and state level risk rates that are constructed from the CFOI 1998-2002 sample, respectively. The risk rates are merged with CPS 1996, 1998, 2000, and 2002 only. The number of workers who face zero risk increases as compared to table 21 and 23. For violent fatal risk, the 1998-2002 CFOI base risk rates is smaller than the 1992-2002 risk rate for total, truck and taxi drivers regardless of the geographic level at which the risk is created. Compared to the 1992-2002 CFOI base risk rate, the mean MSA level violent risk is reduced from 0.83×10^{-4} to 0.46×10^{-4} for the entire sample. The risk rate changes for truck, sales and bus drivers are small, while that of taxi drivers is quite large (5.43×10^{-4} in the 1998-2002 CFOI base risk)

For the non-violent fatal risk, the 1998-2002 CFOI base risk rates are larger than the 1992-2002 CFOI risk rates for all drivers except taxi drivers, regardless the geographic level at which the risk is created. Compared to the 1992-2002 CFOI based risk rate, the mean MSA level non-violent risk is increased from 1.92×10^{-4} to 2.09×10^{-4}, and the mean state level non-violent risk is also increased from 2.92×10^{-4} to 3.16×10^{-4} for the entire sample.

Table 25 shows the summary of injury risks for each driver and event created from the 1992-2002 IFF. For all drivers, the mean violent injury rate is 2.0×10^{-4} and the mean non-violent injury rate is 5.22×10^{-2}. The lowest mean violent injury risk is for truck drivers (1.0×10^{-4}) and the highest mean violent injury risk is for bus drivers (1.2×10^{-3}). The lowest mean non-violent injury risk is for taxi drivers (1.1×10^{-2}) and the highest mean non-violent injury risk is for truck drivers (5.7×10^{-2}). Thirty-percent of drivers face zero violent injury risk, and only 0.1% of drivers face zero non-violent injury risks. None of the truck, sales or bus drivers face zero non-violent injury risks. Table 26 shows the summary of the injury rates created from the 1998-2002 IIF. When using the 1998-2002 IIF, the mean injury risk rates are generally slightly decreased as compared to the 1992-2002 based injury risk. Only the violent injury risks for sales drivers are slightly increased.

Table 1

Definition of Variables and Summary Statistics

	Definition	1996SIPP, all worker-wave (N=166,362) Mean	(SD)
hwage	hourly wage	13.27	(5.90)
risk	fatal risk injury rate by occupation and industry per 10,000 workers	0.55	(0.95)
age	age in years	38.06	(11.47)
ugdeg	1 if individual have bachelor degree or more	0.09	
college	1 if individual attended college	0.33	
hsgrad	1 if individual graduated from high school	0.41	
hispanic	1 if individual has a Hispanic origin	0.13	
blacknh	1 if individual is black and non-Hispanic	0.13	
othrace	1 if individual is non-white, non-black, non-Hispanic	0.04	
female	1 if individual is female	0.45	
workov	1 if individual usually works more than 40 hours	0.19	
union	1 if individual is a union member or covered by union	0.19	
married	1 if individual is married	0.56	
kids18	number of kids under 18 years old	0.79	(1.10)
hipart	1 if individual is provided part of health insurance by employer	0.44	
hifull	1 if individual is provided full health insurance by employer	0.20	
empall	1 if number of employee at all locations > 100	0.56	
empsize	1 if number of employee at worker's location < 25	0.29	
neast	1 if individual lives in Northeastern region	0.16	
midwest	1 if individual lives in Midwestern region	0.26	
west	1 if individual lives in West region	0.18	
south	1 if individual lives in Southern region	0.34	
urban	1 if individual lives in urban area	0.78	
agind	1 if individual works in the agricultural industry	0.02	
constind	1 if individual works in the construction industry	0.07	
tcuind	1 if individual works in the transportation, communications or utility industry	0.06	
trdind	1 if individual works in the wholesale or retail trades industry	0.19	

servind	1 if individual works in the service industry	0.31
manufind	1 if individual works in the manufacturing industry	0.25
pubind	1 if individual works in the public industry	0.10
craftocc	1 if individual has a craft job	0.16
profocc	1 if individual has a professional job	0.11
techocc	1 if individual has a technical job	0.28
servocc	1 if individual has a service job	0.14
farmocc	1 if individual has a farming job	0.05
laborocc	1 if individual has a labor job	0.26

Note. Standard deviations for continuous variables are shown in parenthesis.

Table 2

The Number of Observations Changing Jobs between Major Occupation Groups

After (t) / Before (t-1)	Professional	Craft	Technical	Service	Labor	Farming	Total
Professional	13,854	103	595	196	155	11	14,914
Craft	151	21,438	279	150	716	50	22,784
Technical	752	289	35,388	423	603	31	37,486
Service	243	142	468	16,482	392	38	17,765
Labor	198	833	662	408	33,666	115	35,882
Farming	15	45	53	55	119	2,397	2,684
Total	15,213	22,850	37,445	17,714	35,651	2,642	131,515[a]

[a] The total number in this cell does not match the total number of observations in the sample because for each worker, the first observation is dropped due to the lack of a "previous occupation" data.

Table 3

The Number of Observations Changing Jobs between Major Industry Groups

Before (t-1) \ After (t)	Construc-tion	Agri-culture	TCU	WST	Service	Manufac-turing	Public	Total
Construction	9,380	754	56	144	155	162	25	10,676
Agriculture	51	2,902	14	86	55	34	14	3,156
Transportation /Communication/ Utility (TCU)	52	8	8,785	102	175	84	16	9,222
Wholesales/ Trade (WST)	158	66	133	22,584	834	495	66	24,336
Service	175	43	204	641	38,978	519	147	40,707
Manufacturing	218	48	115	428	546	34,593	46	35,994
Public	21	6	23	35	124	39	7,876	8,124
Total	10,055	3,827	9,330	24,020	40,867	35,926	8,190	132,215[a]

[a] The total number in this cell does not match the total number of observations in the sample because for each worker, the first observation is dropped due to the lack of a “previous occupation” data.

Table 4

The Risk Rate Change in Each Major Occupation Group

Before (t-1) \ After (t)	Professional Mean (sd) [min/max] (unit: 10^{-4})	Craft Mean (sd) [min/max] (unit: 10^{-4})	Technical Mean (sd) [min/max] (unit: 10^{-4})	Service Mean (sd) [min/max] (unit: 10^{-4})	Labor Mean (sd) [min/max] (unit: 10^{-4})	Farming Mean (sd) [min/max] (unit: 10^{-4})
professional	0.00004 (0.037) [-1.04/0.94]	0.722 (1.201) [-0.247/5.301]	-0.049 (0.315) [-0.934/5.584]	0.053 (0.317) [-0.870/0.835]	0.757 (1.264) [-1.378/5.357]	1.115 (0.433) [0.420/1.672]
craft	-0.522 (0.677) [-3.442/0.942]	0.0008 (0.180) [-5.325/4.887]	-0.518 (0.846) [-5.434/5.080]	-0.382 (0.731) [-2.210/1.063]	0.532 (1.467) [-4.851/5.071]	0.461 (2.110) [-4.927/11.372]
technical	0.055 (0.164) [-1.482/0.928]	0.509 (0.795) [-4.867/5.239]	-0.0004 (0.064) [-5.638/5.391]	0.064 (0.359) [-3.819/1.806]	0.753 (1.121) [-4.765/5.402]	1.992 (3.314) [-0.554/12.413]
service	-0.027 (0.334) [-1.814/1.025]	0.534 (0.966) [-1.653/4.774]	-0.051 (0.249) [-1.564/0.737]	0.0001 (0.058) [-1.509/1.521]	0.670 (1.160) [-1.022/4.983]	1.167 (0.699) [-0.374/3.036]
labor	-0.872 (1.266) [-4.270/1.499]	-0.306 (1.347) [-5.013/5.278]	-0.695 (1.058) [-5.640/1.053]	-0.547 (1.111) [-4.534/1.185]	0.004 (0.358) [-5.072/5.500]	0.356 (2.342) [-3.302/12.286]
farming	-0.969 (1.583) [-1.672/0.185]	-0.799 (2.514) [-10.746/4.244]	-1.302 (1.721) [-12.336/2.177]	-1.460 (2.178) [-12.182/0.403]	0.239 (2.117) [-11.505/3.809]	0.002 (0.275) (-2.646/10.695]

Table 5

The Risk Rate Change in Each Major Industry Group

After (t) / Before (t-1)	Construction Mean (sd) [min/max] (unit: 10^{-4})	Agriculture Mean (sd) [min/max] (unit: 10^{-4})	TCU Mean (sd) [min/max] (unit: 10^{-4})	Wholesales/ Trade Mean (sd) [min/max] (unit: 10^{-4})	Service Mean (sd) [min/max] (unit: 10^{-4})	Manufacturing Mean (sd) [min/max] (unit: 10^{-4})	Public Mean (sd) [min/max] (unit: 10^{-4})
Construction	-0.002 (0.391) [-4.826/4.411]	0. 335 (1.909) [-2.721/5.091]	0.067 (2.044) [-3.612/4.317]	-1.500 (1.092) [-4.436/0.590]	-1.667 (1.452) [-5.162/1.917]	-1.986 (1.337) [-5.072/1.724]	-2.022 (1.078) [-3.713/-0.059]
Agriculture	0.290 (1.831) [-3.765/2.622]	0.002 (0.382) [-4.072/5.402]	0.425 (1.553) [-1.690/2.597]	-1.551 (1.135) [-5.434/0.111]	-1.273 (0.937) [-4.327/2.235]	-1.361 (2.484) [-5.325/10.695]	-1.610 (1.665) [-5.367/0.453]
Transportation /Communication/ Utility (TCU)	0.092 (2.097) [-3.769/4.169]	0.240 (2.302) [-2.235/4.253]	0.008 (0.388) [-5.640/5.366]	-1.128 (1.667) [-5.417/1.120]	-0.886 (1.496) [-5.638/2.102]	-1.141 (2.159) [-4.867/12.125]	-0.851 (1.511) [-3.483/0.852]
Wholesales/ Trade	1.471 (1.192) [-0.826/4.943]	1.241 (0.996) [-0.265/4.672]	0.692 (1.508) [-1.361/5.391]	-0.0003 (0.111) [-1.410/1.563]	-0.089 (0.395) [-1.379/2.728]	-0.050 (0.578) [-1.474/2.847]	0.1003 (0.437) [-0.960/0.804]
Service	1.632 (1.373) [-1.364/4.983]	1.546 (1.020) [0.035/5.301]	0.796 (1.485) [-2.798/5.584]	0.077 (0.413) [-2.715/1.399]	0.003 (0.094) [-2.949/2.949]	0.087 (0.900) [-2.194/12.413]	0.134 (0.471) [-2.900/1.366]
Manufacturing	1.815 (1.844) [-10.746/5.094]	1.908 (1.175) [-0.125/5.278]	1.260 (1.530) [-0.946/5.500]	0.061 (0.771) [-11.505/1.578]	-0.060 (0.695) [-12.182/2.383]	0.0006 (0.201) [-12.336/11.831]	0.195 (0.456) [-0.765/0.860]
Public	1.752 (1.226) [-0.424/3.713]	0.700 (0.746) [-0.820/1.650]	1.454 (1.781) [-0.370/4.344]	-0.049 (0.429) [-0.804/0.901]	-0.175 (0.377) [-0.891/1.251]	-0.079 (0.461) [-1.007/0.813]	-0.0002 (0.083) [-1.399/1.382]

Table 6

The Summary of Risk Changes in Different Groups of Workers

	Total	Mean[a] (sd) [min/max] (unit:10^{-4})	# of workers whose risk change is not zero (%in Total)	Mean (sd) [min/max] (unit:10^{-4})	Median (unit:10^{-4})
Male workers	72,212	0.0039 (0.46) [-12.33/12.41]	7,646 (10%)	0.0376 (1.43) [-12.33/12.41]	0.0009
Female workers	59,303	0.0002 (0.15) [-5.64/4.95]	6,087 (10%)	0.0020 (0.49) [-5.64/4.95]	0.0001
Single workers	55,029	0.0031 (0.39) [-12.18/12.41]	6,821 (12%)	0.0226 (1.09) [-12.18/12.41]	0.0008
Married workers	76,486	0.0016 (0.33) [-12.33/12.12]	6256 (8%)	0.0193 (1.15) [-12.33/12.12]	0.0010
Workers who have kids age under 18	22,507	0.0013 (0.33) [-10.74/10.83]	2,294 (10%)	0.0135 (1.06) [-10.74/10.83]	0.0021
Workers who do not have kids age under 18	72,380	0.0009 (1.13) [-12.33/12.41]	7,131 (9%)	0.0099 (1.13) [-12.33/12.41]	-0.0008
Workers in union	23,471	0.0024 (0.25) [-10.74/4.96]	984 (4%)	0.0588 (1.25) [-10.74/4.96]	0.0122
Workers in non-union	100,831	0.0018 (0.36) [-12.33/12.41]	11,335 (11%)	0.0167 (1.08) [-12.33/12.41]	0.0001

[a] The median value of all cases is zero.

Table 7

The Summary of Risk Changes in Different Groups of Workers Who Changes Demographic Status

	Total	Mean[a] (sd) [min/max] (unit:10^{-4})	# of workers whose risk change is not zero (%in Total)	Mean (sd) [min/max] (unit:10^{-4})	Median (unit:10^{-4})
Single to married	1,503	0.0037 (0.50) [-5.32/5.09]	429 (28%)	0.0131 (0.94) [-5.32/5.09]	-0.0113
Married to single	805	0.0236 (0.63) [-3.98/11.37)	227 (28%)	0.0839 (1.19) [-3.98/11.37]	-0.0008
No kids under 18 to some kids under 18	2,358	0.0099 (0.54) [-4.99/5.30]	552 (23%)	0.0425 (1.13) [-4.99/5.30]	0.0070
Some kids under 18 to no kids under 18	3,341	0.0189 (0.57) [-5.32/11.37]	698 (20%)	0.0908 (1.26) [-5.32/11.37]	0.0018
Union to non-union	3,660	-0.0147 (0.50) [-5.41/4.20]	652 (17%)	-0.0828 (1.18) [-5.41/4.20)	-0.0161
Non union to union	3,553	0.0301 (0.64) [-11.58/10.83]	762 (21%)	0.1405 (1.38) [-11.58/10.83]	0.0124

[a] The median value of all cases is zero.

Table 8

Variable Definition, Source and Summary Statistics (n=19,371[a])

Variables	*Description*	*Mean (SD)*	*Data source*
wage	weekly wage (adjusted to $2004)	683 (336)	CPS (94, 96, 98, 00, 02)
age	age in years	40.69 (11.60)	
ugdeg	1 if individual has a four year college degree	0.04	
college	1 if individual attended college	0.23	
hsgrad	1 if individual graduated from high school	0.52	
hispanic	1 if individual has a Hispanic origin	0.10	
blacknh	1 if individual is black and non-Hispanic	0.12	
othrace	1 if individual is not black, white or Hispanic	0.02	
uscit	1 if individual is U.S. citizen	0.93	
female	1 if individual is female	0.06	
salary	1 if individual works for salary	0.36	
workov	1 if individual usually works more than 40 hours	0.50	
union	1 if individual is a union member or covered by union	0.23	
married	1 if individual is married	0.66	
central	1 if individual lives in central city	0.21	
truck	1 if individual is truck driver	0.83	
bus	1 if individual is bus driver	0.07	
taxi	1 if individual is taxi driver	0.05	
year dummy	1 if individual's data comes from corresponding year		
MSA dummy	1 if individual lives in a corresponding MSA		
State dummy	1 if individual lives in a corresponding state		
Region dummy	1 if individual lives in a corresponding region		
nonfatal_vrisk	occupational nonfatal violent assault risk	2.18×10^{-4} (5.36×10^{-4})	estimated by authors from IIF (92-03) and OES (98-03)
nonfatal_other	occupational nonfatal non-violent injury risk	5.25×10^{-2} (1.63×10^{-2})	
unemp	annual unemployment rate of MSA where individual lives	5.10 (1.95)	Local Area Unemployment Statistics (94, 96, 98, 00, 02)
whole_sales	per capita sales in whole industry in MSA where individual lives ($1,000)	17.58 (9.95)	1997 Economic Census
retail_sales	per capita sales in retail industry in MSA where individual lives ($1,000)	9.38 (1.91)	
trans_sales	per capita sales in transportation industry in MSA where individual lives ($1,000)	1.39 (1.04)	
ent_sales	per capita sales in entertainment industry in MSA where individual lives ($1,000)	0.08 (0.05)	
msavmtp	per capita vehicle miles traveled in MSA where individual lives	8,315 (2,901)	Bluestone (forthcoming)

Note. The variables unemp, whole_sales, retail_sales, trans_sales, ent_sales, msavmtp, msavmta, ppmsa contain less observations due to the missing values.

Table 9

Summary of Fatal Incidences by Occupation and Event (1992-2002 CFOI)

		Truck	Sales	Bus	Taxi	Total
Violent	Total	232	101	25	668	1,026
	(in an MSA)	(195)	(88)	-[a]	(612)	-[a]
	[MSA % of total]	[84]	[87]	-[a]	[91]	-[a]
Non-violent						
Transportation	Total	7,176	335	150	245	7,906
	(in an MSA)	(3,776)	(214)	(97)	(193)	(4,280)
	[MSA % of total]	[52]	[63]	[64]	[78]	[54]
Others	Total	1,362	29	16	16	1,423
	(in an MSA)	(862)	(20)	-[a]	-[a]	(907)
	[MSA % of total]	[63]	[68]	-[a]	-[a]	[63]

[a] The number is suppressed for reasons of confidentiality.

Table 10

MSAs with Top Five Number of Deaths by Occupation and Event (1992-2002 CFOI)

Violent fatal injury	# of deaths (risk rate)	Non-violent fatal injury	# of deaths (risk rate)
Truck driver			
Los Angels-Long Beach, CA PMSA	16 (0.22×10^{-4})	Riverside-San Fernardino, CA PMSA	140 (4.93×10^{-4})
Dallas, TX PMSA	9 (0.19×10^{-4})	Los Angels-Long Beach, CA PMSA	126 (1.79×10^{-4})
Chicago, IL PMSA	7 (0.19×10^{-4})	Chicago, IL PMSA	113 (1.38×10^{-4})
Washington D.C, PMSA	7 (0.19×10^{-4})	Atlanta, GA MSA	94 (1.91×10^{-4})
Miami FL, PMSA	6 (0.35×10^{-4})	Houston, TX PMSA	94 (2.29×10^{-4})
Sales driver			
Philadelphia, PA-NJ PMSA	8 (1.43×10^{-4})	Philadelphia, PA-NJ PMSA	11 (1.97×10^{-4})
Dallas, TX PMSA	6 (1.32×10^{-4})	Atlanta, GA MSA	10 (1.42×10^{-4})
Atlanta, GA MSA	6 (0.85×10^{-4})	Chicago, IL PMSA	7 (0.81×10^{-4})
Detroit, MI PMSA	5 (0.90×10^{-4})	Houston, TX PMSA	7 (0.85×10^{-4})
Los Angels-Long Beach, CA PMSA	-[a]	Los Angels-Long Beach, CA PMSA	7 (0.51×10^{-4})
Taxi driver			
New York, NY PMSA	186 (32.26×10^{-4})	New York, NY PMSA	25 (4.3×10^{-4})
Chicago, IL PMSA	29 (9.39×10^{-4})	Chicago, IL PMSA	9 (2.91×10^{-4})
Atlanta, GA MSA	24 (13.41×10^{-4})	Washington D.C, PMSA	7 (2.43×10^{-4})
Washington D.C, PMSA	24 (8.33×10^{-4})	Los Angels-Long Beach, CA PMSA	7 (1.42×10^{-4})
Los Angels-Long Beach, CA PMSA	20 (4.36×10^{-4})	Philadelphia, PA-NJ PMSA	6 (2.36×10^{-4})

[a] The number is suppressed for reasons of confidentiality

Table 11

States with Top Five Number of Death Incidences by Occupation and Event (1992-2002 CFOI)

Violent fatal injury	# of deaths (risk rate)	Non-violent fatal injury	# of deaths (risk rate)
	<u>Truck driver</u>		
California	36 (0.14×10^{-4})	California	821 (3.20×10^{-4})
Texas	27 (0.13×10^{-4})	Texas	777 (3.93×10^{-4})
Florida	15 (0.11×10^{-4})	Florida	418 (3.08×10^{-4})
Illinois	10 (0.07×10^{-4})	North Carolina	351 (4.17×10^{-4})
New York	10 (0.08×10^{-4})	Pennsylvania	332 (2.49×10^{-4})
	<u>Sales driver</u>		
Pennsylvania	10 (0.55×10^{-4})	Texas	34 (1.04×10^{-4})
Texas	9 (0.27×10^{-4})	Georgia	30 (2.27×10^{-4})
Georgia	8 (0.60×10^{-4})	Missouri	21 (2.24×10^{-4})
New York	6 (0.38×10^{-4})	Ohio	20 (0.93×10^{-4})
Florida	6 (0.21×10^{-4})	Pennsylvania	20 (1.11×10^{-4})
	<u>Taxi driver</u>		
New York	200 (18.88×10^{-4})	New York	40 (3.77×10^{-4})
California	48 (3.23×10^{-4})	California	23 (1.54×10^{-4})
Florida	36 (3.73×10^{-4})	Florida	22 (2.28×10^{-4})
Illinois	33 (6.90×10^{-4})	New Jersey	15 (2.32×10^{-4})
Georgia	33 (10.07×10^{-4})	Illinois	13 (0.20×10^{-4})

Table 12

National Trend of Employment Growth in Driving Occupations

Occupation	Average employment 1992-1997 (thousands)	Average employment 1998-2003 (thousands)	Percent change between 1992-1997 average and 1998-2003 average
Truck	1217.66	1367.23	12% increase
Taxi	31.35	32.4	3% increase
Bus	21.85	22.86	4% increase

Table 13

Summary of Fatal Incidence by Occupation and Event in the 1998-2002 CFOI (in parenthesis, the proportion of deaths in 1992-2002 is reported)

	Truck	Sales	Bus	Taxi	Total
Violent					
Total	85 (36%)	43 (42%)	12 (48%)	208 (31%)	348 (33%)
(in an MSA)	74 (37%)	38 (43%)	-[a]	190 (31%)	-[a]
Non-violent					
Total	4,125 (48%)	194 (53%)	102 (61%)	130 (49%)	4,551 (48%)
(in an MSA)	2,259 (48%)	124 (52%)	67 (61%)	101 (49%)	2,551 (49%)

[a] The number is suppressed for reasons of confidentiality

Table 14

Mean and Standard Deviation of Fatal Risk Rate by Occupation and Event (1992-2002 CFOI) (standard deviation reported in parentheses)

	MSA level		State level	
	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)
Total	0.97 (3.97)	1.74 (3.72)	1.42 (4.20)	1.99 (1.64)
Truck	0.06 (0.15)	2.62 (2.10)	0.08 (0.15)	3.27 (1.68)
Sales	0.24 (1.00)	1.02 (2.19)	0.27 (0.27)	0.98 (0.70)
Bus	0.13 (0.85)	0.95 (2.95)	0.16 (0.43)	1.45 (1.59)
Taxi	3.52 (7.36)	2.17 (6.09)	5.34 (7.30)	2.14 (1.39)

Table 15

Mean and Standard Deviation of MSA Level Fatal Risk Rate by Occupation and Event (1992-2002 CFOI) After Dropping MSAs With Less Than 100 Employees (standard deviation reported in parentheses)

	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)
Total	0.75 (2.82)	1.60 (2.29)
Truck	0.06 (0.15)	2.62 (2.10)
Sales	0.23 (0.82)	0.97 (1.91)
Bus	0.11 (0.51)	0.68 (1.44)
Taxi	3.75 (5.88)	1.68 (3.10)

Table 16

Mean and Standard Deviation of Fatal Risk Rate by Occupation and Event Based on 1998-2002 CFOI (standard deviation reported in parentheses)

	MSA level		State level	
	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)
Total	0.75 (4.39)	1.95 (5.09)	1.00 (3.11)	1.91 (1.91)
Truck	0.05 (0.22)	2.79 (2.46)	0.04 (0.05)	3.32 (2.19)
Sales	0.19 (1.11)	1.22 (3.63)	0.26 (0.45)	1.09 (1.10)
Bus	0.14 (1.64)	1.43 (5.20)	0.19 (0.83)	1.43 (1.76)
Taxi	2.66 (8.40)	2.25 (7.80)	3.47 (5.47)	1.79 (1.65)

Table 17

Mean and Standard Deviation of MSA Level Fatal Risk Rate by Occupation and Event After Dropping MSAs With Less Than 100 Employees Based on 1998-2002 CFOI (standard deviation reported in parentheses)

	Violent assault (in 10,000)	Non-violent fatal risk (in 10,000)
Total	0.48 (2.33)	1.78 (3.20)
Truck	0.05 (0.22)	2.79 (2.46)
Sales	0.21 (1.15)	1.03 (2.53)
Bus	0.05 (0.31)	0.92 (2.55)
Taxi	2.29 (5.06)	2.03 (5.09)

Table 18

Missing Injury Data by State[a]

Name of State	Years missing
Colorado	1992-2002
District of Columbia	1992-2002
Idaho	1992-2002
Illinois	1992-1997
Maine	1995
Maryland	1995
Mississippi	1995-2002
Missouri	1995
Montana	1995
New Hampshire	1992-2002
New Jersey	1992
North Carolina	1996
North Dakota	1992-2002
Ohio	1992-2002
Oregon	1995
Pennsylvania	1992, 1995-2002
South Carolina	1995
Vermont	1995-1996
Virginia	1995
West Virginia	1992-1997
Wyoming	1995-2001

[a] State not listed had no missing data.

Table 19

State Level Non-fatal Injury Rates by Occupation Based on the IFF 1992-2002 (standard deviation reported in parentheses)

	Truck drivers	Sales drivers	Bus drivers	Taxi drivers
Assault injury	0.0001 (0.0001)	0.0002 (0.0009)	0.0004 (0.0008)	0.00007 (0.0003)
Transportation related injury	0.007 (0.001)	0.003 (0.002)	0.004 (0.003)	0.002 (0.003)
Total injury	0.0581 (0.011)	0.046 (0.018)	0.019 (0.015)	0.007 (0.007)

Table 20

State Level Non-fatal Injury Rates by Occupation Based on the IFF 1998-2002 (standard deviation reported in parentheses)

	Truck drivers	Sales drivers	Bus drivers	Taxi drivers
Assault injury	0.00005 (0.0001)	0.0003 (0.0009)	0.0003 (0.0007)	0.00006 (0.0004)
Transportation related injury	0.007 (0.001)	0.004 (0.003)	0.004 (0.003)	0.001 (0.003)
Total injury	0.055 (0.012)	0.041 (0.014)	0.019 (0.014)	0.006 (0.008)

Table 21

Summary of MSA Level Fatal Risk Rates for the Sample of Workers (Risk Created Based on 1992-2002 CFOI; standard deviation reported in parentheses)

	N	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)	# of workers with zero violent/non-violent risk
Total	12,892	0.83 (4.19)	1.92 (1.63)	5,408/1037
Truck	10,293	0.08 (0.11)	2.07 (2.46)	4,036/ 189
Sales	602	0.30 (0.57)	0.70 (3.63)	366/231
Bus	1,115	0.18 (1.64)	0.87 (1.39)	880/440
Taxi	882	10.67 (12.33)	2.38 (3.17)	126/177

Table 22

Summary of State Level Fatal Risk Rates for the Sample of Workers (Risk Created Based on 1992-2002 CFOI; standard deviation reported in parentheses)

	N	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)	# of workers with zero violent/non-violent risk
Total	19,371	0.52 (2.71)	2.92 (1.67)	3,505/200
Truck	16,085	0.07 (0.08)	3.23 (1.56)	2627/ 0
Sales	895	0.26 (0.23)	0.91 (0.58)	195/33
Bus	1,385	0.11 (0.23)	0.92 (1.06)	629/130
Taxi	1,006	8.38 (8.73)	2.39 (1.32)	54/37

Table 23

Summary of MSA Level Fatal Risk Rates for the Sample of Workers (Risk Created Based on 1998-2002 CFOI; standard deviation reported in parentheses)

	N	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)	# of workers with zero violent/non-violent risk
Total	10,479	0.46 (1.91)	2.09 (1.84)	6,215/1155
Truck	8,369	0.06 (0.11)	2.23 (1.60)	4,862/ 254
Sales	483	0.36 (1.04)	0.76 (1.48)	358/283
Bus	906	0.17 (0.96)	1.38 (2.78)	823/392
Taxi	721	5.43 (4.98)	2.37 (2.41)	172/226

Table 24

Summary of State Level Fatal Risk Rates for the Sample of Workers (Risk Created Based on 1998-2002 CFOI; standard deviation reported in parentheses)

	N	Violent fatal risk (in 10,000)	Non-violent fatal risk (in 10,000)	# of workers with zero violent/non-violent risk
Total	15,447	0.32 (1.58)	3.16 (1.98)	4,600/348
Truck	12,829	0.05 (0.05)	3.48 (1.93)	3,580/ 0
Sales	689	0.27 (0.38)	1.15 (0.96)	265/70
Bus	1,118	0.13 (0.47)	1.27 (1.23)	658/189
Taxi	811	4.90 (4.99)	2.38 (1.35)	97/89

Table 25

Summary of State Level Injury Rates for the Sample of Workers (Injury Risk Created Based on 1992-2002 IFF; standard deviation reported in parentheses)

	N	Violent injury risk	Non-violent injury risk	# of workers with zero violent/non-violent injury risk
Total	17,091	0.0002 (0.0005)	0.0525 (0.0163)	5,142/22
Truck	14,164	0.0001 (0.0001)	0.0577 (0.0093)	2,215/ 0
Sales	772	0.0002 (0.0007)	0.0466 (0.0148)	772/0
Bus	1,230	0.0012 (0.0014)	0.0279 (0.0179)	1,230/0
Taxi	925	0.0002 (0.0004)	0.0115 (0.0078)	925/22

Table 26

Summary of State Level Injury Rates for the Sample of Workers (Injury Risk Created Based on 1998-2002 IFF; standard deviation reported in parenthesis)

	N	Violent injury risk	Non-violent injury risk	# of workers with zero violent/non-violent injury risk
Total	12,754	0.0001 (0.0005)	0.0485 (0.0159)	7,747/105
Truck	10,529	0.00005 (0.0001)	0.0534 (0.0103)	6,286/ 0
Sales	561	0.0003 (0.0008)	0.0409 (0.0139)	436/70
Bus	945	0.0010 (0.0012)	0.0271 (0.0154)	344/189
Taxi	719	0.0001 (0.0006)	0.0111 (0.0090)	681/105

Chapter IV

Panel Data Analysis of Hedonic Wage Model

This study uses panel data models to identify the endogeneity bias in the cross-sectional HW model and to estimate a consistent wage/risk premium. In addition to employing panel models to control for time-invariant omitted variables, we also test the strict exogeneity assumption to assure the consistency of panel estimators. To my knowledge, this is the first study to test the use of panel estimators in estimating the HW model. The labor market data comes from the large continuous national panel, the Survey of Income and Program Participation (SIPP) panel administered by U.S. Census Bureau. As discussed in chapter 3, the SIPP contains rich information on workers' wage and demographic characteristics. The occupational fatal risk comes from Scotton (2000). The risk rate varies by occupation and industry group.

Endogeneity in Hedonic Models.

The standard HW model estimates the following equation by an ordinary least square (OLS) model:

$$y_i = \beta r_i + X_i \gamma + \mu_i \tag{15}$$

where y_i is a wage level (often natural log form), r_i is an occupational fatal risk level, X_i is a vector of determinants of wages (e.g., age, educational attainment, race, sex, including a constant), and μ_i is an error term for individual i.

The coefficient of risk variable β represents the risk premium on wages. The estimated β is unbiased only if r_i is an exogenous variable. If the risk variable is endogeneous such that the $\text{cov}(r_i, \mu_i | X_i) \neq 0$, then the estimated β is biased and inconsistent. The non-zero covariance between the risk variable and the error term may arise from omitted variables, measurement error associated with the risk variable, or simultaneity between wage and the risk variable.

In the HW literature, many researchers show concern for this endogeneity problem in the cross-section HW models, especially endogeneity related to omitted variables or measurement error associated with risk variables. Potential unobservable characteristics which could influence wages and would vary with a worker's wage-risk tradeoff include heterogeneous workers' characteristics such as risk preference or productivity under non-safe environment, and working conditions such as the level of physical exertion involved and the risk of job-related injuries (Viscusi & Aldy, 2003).

Measurement error associated with the risk variable can be one of two types. One type is measurement error between the estimated objective risk level and the actual individual worker's risk level. The other type is measurement error between the actual individual worker's risk level and worker's perceived risk level (see discussion about measurement error in chapter 2). The first type of measurement error has not been considered in the HW literature until recently. Most of the HW studies in the U.S. assign the industry average risk level to each worker to estimate the HW equation (Viscusi and Aldy 2003). This practice assumes that different occupations in a same industry such as the secretary and construction worker at construction industry face the same risk level, which is apparently not true.

The recent renovation in the occupational risk data enables researchers to construct more flexible format of the risk variable, such as the risk rate varies by occupation and industry group as used in this study.[42] It is still possible that there is measurement error between the estimated risk level and the actual individual worker's risk level, but error of this type should be smaller when we use the occupation-industry risk rate as compared to the industry average risk rates.

There is no satisfactory argument about the second type of measurement error (McConnell 2006). However, past studies report important disparities between subjective risk measures (the risk a person believes he/she faces) and objective risk measures for different types of risks (Benjamin, 2001; Lichtenstein, Slovic, Fischoff, Layman, & Combs, 1978), and thus this type of error may bias the risk estimator significantly.

Several empirical studies attempt to correct this endogeneity bias in the HW model using an instrumental variables (IV) approach. Siebert and Wei (1994) use a standard 2SLS model. Garen (1988) use a general case of the 2SLS approach which includes the interaction term between the estimated residual from the first stage equation and the risk variable in addition to the risk variable itself (Card, 1999). Gunderson and Hyatt (2001) and Arabsheibani and Marin (2001) follow Garen's approach using labor market data in Canada and in the United Kingdom, respectively.

Previous studies generally find a substantial increase of the risk premium when an IV estimation strategy is used. For example, Siebert and Wei (1994) find a 1.5 to 2.5-fold increase, Gunderson and Hyatt (2001) find a five-fold increase, and Arabsheibani and Marin (2001) find up to a 10-fold increase in the estimated risk premium. However, results of these studies are often quite sensitive to the model's specification. Arabsheibani and Marin (2001) find that a slight change in the sample composition and in the model specification dramatically changes the magnitude of their estimated risk

[42] See Scotton (2000) and Viscusi (2004) for the studies which use occupation-industry risk rates.

premium in the Garen model. They argue that the instability of their risk premium estimates are likely due to the poor fit of the first stage risk equation.

There are a few studies that use panel data to correct the time invariant unobserved heterogeneity problem in the HW equation, but the results of these are mixed. Brown (1980) estimates a HW model using a fixed effect (FE) model with panel data of the National Longitudinal Survey of Youth (NLSY). Although he uses an actuarial risk measure, which reflects not only occupational causes of death, but all causes of death, he finds a statistically significant positive estimate of the risk premium. Black et al. (2003) also use NLSY for their analysis, but could not find a statistically significant result, likely due to a large measurement error bias associated with their measure of workplace risk. Kniesner et al. (2005) use a first difference (FD) model on panel data from the Panel Survey of Income Dynamics (PSID), incorporating a fairly disaggregated occupational risk measure. They find a significant decrease in risk premium once they control for the time-invariant heterogeneities.

Although Kniesner et al. (2005) provide the most reliable risk premium estimate among previous panel studies because they use the most disaggregated occupational risk data, there still remains concern. As Griliches and Hausman (1986) indicate, it is common to observe lower or insignificant estimates when one applies a FE or FD model as these models exacerbate measurement error bias. Thus smaller or insignificant coefficient estimates may be caused by measurement error bias, and not necessarily due to the correction of unobserved heterogeneity. In addition, Kniesner et al. (2005) assume that time-varying unobserved factors that could be correlated with risk are negligible, but it may not be the case. For instance, on the job training in a previous job might affect the worker skill related to the next job, although on-the-job training is often unobserved to the researcher

VSL is an important component of policy analysis, and careful examination of this potential bias in the panel data analysis should be conducted. The panel data analysis and IV approach have different advantages and disadvantages. Panel data provides an excellent way to control for time invariant unobserved heterogeneity, but panel models are sensitive to measurement error bias. The IV approach can control for more endogenous factors, but requires additional information.

In this study, we estimate panel models to assess the potential bias due to endogeneity problems in cross-section HW models. We test the consistency of panel estimators by employing simple statistical tests, as well as by combining the panel data approach and the IV approach to control for time-invariant and time-varying worker heterogeneity, measurement error bias, and simultaneity between wage and risk variable. The next section first discusses the theory underlying the FD and FE models, followed by a discussion of the strict exogeneity assumption to obtain consistent panel estimators. We then present the estimation results for the pooled cross-section OLS, FD and FE models. After discussing these results, we examine potential violations of the strict exogeneity assumption in panel models from the simple statistical test and from two stage panel models. We also conduct sensitivity analysis on our results. Lastly we present the conclusions of the analyses.

Basic FD and FE Model

Assume the wages in period t, $y_{i,t}$, are determined as follows:

$$y_{i,t} = \beta r_{i,t} + X_{i,t}\gamma + Z_i\delta + \mu_{i,t}\,, \qquad (16)$$

where $y_{i,t}$, $r_{i,t}$ and $X_{i,t}$ are defined as in the equation (15), Z_i is a vector of unobserved time-invariant factors, $\mu_{i,t}$ is an error term which may contain a vector of unobserved time-varying factors or measurement error associated with the risk variable, and $t = \{1,2,...,T\}$.

A FD model with T periods implies the following estimating equation:

$$\Delta y_{i,t} = \beta \Delta r_{i,t} + \Delta X_{i,t}\gamma + \Delta \mu_{i,t}, \tag{17}$$

where,

$\Delta y_{i,t} = y_{i,t} - y_{i,t-1}$, $\Delta r_{i,t} = r_t - r_{t-1}$, $\Delta X_{i,t} = X_{i,t} - X_{i,t-1}$, and $\Delta \mu_{i,t} = \mu_{i,t} - \mu_{i,t-1}$.

For a FE model with T periods, the estimating equation becomes:

$$\ddot{y}_{i,t} = \beta \ddot{r}_{i,t} + \ddot{X}_{i,t}\gamma + \ddot{\mu}_{i,t} \tag{18}$$

where $\ddot{y}_{i,t} = y_{i,t} - \bar{y}_i$, $\ddot{r}_{i,t} = r_{i,t} - \bar{r}_i$, $\ddot{X}_{i,t} = X_{i,t} - \bar{X}_i$, and $\ddot{\mu}_{i,t} = \mu_{i,t} - \bar{\mu}_i$, and where

$$\bar{y}_i = \frac{\sum_{t=1}^{T} y_{i,t}}{T}, \bar{r}_i = \frac{\sum_{t=1}^{T} r_{i,t}}{T}, \bar{X}_i = \frac{\sum_{t=1}^{T} X_{i,t}}{T} \text{ and } \bar{\mu}_i = \frac{\sum_{t=1}^{T} \mu_{i,t}}{T}.$$

Note that in both equations, the unobserved time-invariant heterogeneity *Z* is perfectly controlled, while remaining unobserved factors may be still a problem. We estimate equation 16 and 17 to obtain the FD and FE estimators.

Endogeneity in Panel Models

The important underlying assumption that panel models must satisfy to obtain consistent estimators is the strict exogeneity assumption.. The strict exogeneity assumption requires that the error terms are uncorrelated with *r*, *X* or *Z* in any time period as expressed in equation (19):

$$E(\mu_{i,t} \mid r_{i,s}, X_{i,s}, Z_i) = 0, \text{ for } \forall \text{ t,s} \in \text{T}. \tag{19}$$

The violation of strict exogeneity may come from contemporaneous correlation between the risk variable and the error term, which can be caused by time-varying omitted variables, measurement error associated with risk variable, or simultaneity between wages and the risk variable (Wooldridge, 2001). The violation of strict exogeneity assumption leads to inconsistent panel estimators. In this study, we assume that a vector of *X* satisfies the strict exogeneity assumption. Thus we only focus on *r* as a potential endogeneous variable.

According to Wooldridge (2001), one can test for the strict exogeneity assumption of variable of interest, using the FE model with t>2 by testing a H_0: $\xi = 0$ by estimating on the FE model assuming the HW model is

$$y_t = \beta r_t + X_t \gamma + W_{t+1}\xi + Z\delta + \mu_t \tag{20}$$

where *W* is the subset of *X*. In our case, *W* consists only of the risk variable. If there is a violation of strict exogeneity, one msut find external instruments to obtain consistent estimators.

For the FD model, the outside instruments, *Q*, should satisfy following conditions:

$$\mathrm{cov}(\Delta u_{i,t}, Q_{i,t}) = 0 \text{ and } \mathrm{cov}(\Delta r_{i,t}, Q_{i,t}) \neq 0 \tag{21}$$

And for the FE model, the outside instrument should satisfy:

$$\mathrm{cov}(\ddot{u}_{i,t}, Q_{i,t}) = 0 \text{ and } \mathrm{cov}(\ddot{r}_{i,t}, Q_{i,t}) \neq 0 \tag{22}$$

The conditions above imply that the variables *Q* are expected to influence the choice of risk but not wages received. More specifically, *Q* is expected to be correlated with changes in risks but not changes in wages in the FD models. And for the FE model, *Q* is expected to influence time-demeaned risks but not time-demeaned wages.

In the next section, we present the OLS, FD and FE estimation results, followed by the test results for the strict exogeneity assumption. We also examine the potential endogeneity problem from the contemporaneous correlation using external instruments, assuming there are time-varying omitted variables, measurement error associated with risk variable or simultaneity between wage and risk variables.

Results

Table 27 shows the estimation results with OLS, FD and FE models. Data are discussed in chapter 3. The dependent variable is the log of gross hourly wage level, and independent variables include the occupational risk level, job characteristics, worker's characteristics, and regional variables. Most of explanatory variables are statistically significant, and results are generally consistent with findings in previous studies using similar risk measures such as Scotton (2000) and Viscusi (2004). Age and education level are positively correlated with wages. Hispanics and African Americans earn less than whites, and females earn less than males. Workers who belong to a union receive higher wages than non-union workers and so do married workers compared to single workers.

The coefficient for the risk variable from the cross-section HW model is 0.0167 and statistically significant at the 1% level. The cross section HW model allows the correlation among observations of a same worker. The estimated VSL from this cross-section HW model is $4.6 million.[43] This value is in a range of the average value of previous VSL estimates (Kochi et al., 2006; Viscusi, 1992). This VSL estimate is significantly smaller than the estimate based on the similar risk measure by Kniesner et

[43] The VSL is estimated as follows: VSL=coefficient of risk variable ×hourly wage× 40 ×52 (weeks) ×10,000, where 10,000 is the unit of fatal risk.

al. (2005), which is $17.7 million.[44] The difference between the estimates likely comes from the different model specification as well as the difference in average wage level of the sample. Kniesner et al. (2005) do not include industry dummy variables due to the concern of multicolinearity between risk variables and industry dummy variables. In addition, Kniesner et al. (2005) do not include firm side variables included in our model. The VSL estimate also depends on the average wage level. In our sample, the average wage is $13.27 while the sample of Kniesner et al. (2005) has the average wage of $21.04.

When we omit industry dummy variables from our model, the estimated coefficient of the risk variable increases to 0.0374, which also increases the VSL to $10.3 million. This estimate is still less than the estimates of Kniesner et al. (2005). When we omit firm specific variables, *hipart*, *hifull*, *empsiz*, *empall*, in addition to omitting industry dummy variables, the estimated coefficient of the risk variable decreases to 0.0264, which is a VSL of $7.2 million. Both these estimates are higher than the VSL obtained from our pooled cross-section OLS model, which is $4.6 million. This indicates that the excluding industry dummy variables and firm-side variables likely overestimates the wage/risk premium.

As shown in table 27, all industry dummy variables and firm variables are significant factors to determine wages and they also affect the coefficient of risk. Including these variables does not reduce the significance of risk coefficients, which indicates that the multicollinearity is not likely the issue in our model. Since these variables are correlated with risk variable, excluding these variables would bias the results and thus we keep these variables in our models.

[44] Estimate based on 11-year average risk rates.

Remaining differences in the VSL estimates between our study and Kniesner et al (2005) are likely from the difference in the wage level across sample of workers. As mentioned earlier, the average wage (in our sample) is $13.27 while that of Kniesner et al.'s (2005) sample is $21.04. Due to the over-sampling of low-income population in the SIPP program, the average wage level of the sample in the SIPP is less than national average (see chapter 3). The average wage level of Kniesner et al.'s (2005) sample is higher than the national average, which we would expect to make the divergence between our estimate and their estimate larger if risk is a normal good.

Our cross section HW result indicates that workers in the U.S. labor market receive a significant wage premium for accepting higher levels of occupational risk. However, when we apply the FE model and the FD model, the coefficient of risk variable dramatically reduced, yet still significant. The FE and FD models show the coefficient of 0.0094 and 0.0062, respectively, which are significant at the 1% level. The estimated VSL are $2.5 and $1.7 million, respectively. The 95% confidence interval of the VSL based on the FE estimator is $1.7 -$3.4 million and that of the VSL based on the FD estimator is $0.5 -$2.9 million. These results indicate that the unobserved time invariant worker characteristics significantly bias the OLS results upward, and resulted in the pooled OLS overestimating the wage-risk premium. This finding is similar to Kniesner et al. (2005). They also find that panel models significantly reduce the risk coefficient. However, after they use FD models, their estimated VSL becomes $6.7 million (with approximate 95% interval of $2.7-$10.7million), which is still higher than our point estimates, and just barely overlaps with the 95% confidence interval of our FD estimate.[45]

[45] When we omit the industry dummy and firm-side variables, our FE and FD estimator is 0.014 with standard error 0.0013 and 0.011 with standard error 0.0019, respectively. The corresponding VSL

The estimated risk coefficients of our FD and FE models are similar to each other with overlapped 95 % intervals, which may indicate that there is not an important endogeneity issue in the models.[46] However, it is still worthwhile to examine the consistency of estimators using additional methods because the panel estimators of risk variable are not consistent if the risk variable is endogenous in the panel. The following section examines the consistency of FE and FD estimators using two methods. The first method employs the strict exogeneity test illustrated in Wooldridge (2001), and the second method employs the second stage panel models.

First, we examine the consistency of risk estimators in panel models using the strict exogeneity testing method illustrated in Wooldridge (2001). Following Wooldridge (2001), we include one period lead risk variable ($risk_{t+1}$), and re-estimate the FE model, assuming all other explanatory variables are strictly exogeneous. The estimating coefficient of lead risk variable is -0.0004 with standard error 0.0015. The coefficient is not significant, which indicate that the strict exogeneity assumption is not violated in our models. However, when we estimate model without industry variables, where the risk variable is clearly endogenous, the lead risk variable is still not significantly different from zero. This indicates that this strict exogeneity test may not be strong enough.

Next, we employ the two stage panel models to ensure the exogeneity of risk variable. Two types of instruments are explored. The first type of instrumental variables are “outside” the model data and are variables that we expect will influence the choice of risk-level but not wages-received. These are; the monthly income other than wage

---

estimates are $3.8 (95% interval: $3.1-$4.5million) and $3.0 million (95% interval: $2-4million), respectively. These change indicate that the Kniesner et al. (2005) study overestimate the VSL.

[46] Ziliak et al. (1999) noted that if there is no endogeneity in panel models and if the FE model is adjusted for a non-stationarity, then the FD model and the FE model should have a same probability limit when more than two time period are contained in the data.

(*inc_other*), the number of social security recipients in the household (*N_SS*), the monthly income from all financial investments (*inv_all*) and a dummy variable indicating that the reason the employee chose not to have health insurance is a lack of belief in health insurance (*nohi_reason*).[47] These variables are obtained from the 1996 SIPP.

The wage level should be determined according to the worker's productivity. The incomes that are earned through non-wage sources, the number of social security recipients in household, or their lack of belief in health insurance would not likely affect the worker's productivity. On the other hand, the level of total wealth, number of dependents or belief in health insurance may be related to the worker's risk taking behavior.

The second type of instruments adds a variable developed from the risk data itself. This additional variable is the difference between the risk level of individual worker and the average risk of the 3-digit level occupation in which the worker engages (*dif_rocc*). The variable *dif_rocc* is expected to have a strong correlation with the risk variable. The variable *dif_rocc* is a valid instrument only if the worker's deviation from the mean risk level within a same occupation is not correlated with the error term.[48]

Table 28 shows the first-stage regression results with the second type of instrumental variables, and table 29 shows the second-stage regression results from IV-

---

[47] There are several categories of reasons employees chose not to be covered by health insurance in the 1996 SIPP, and they are: health insurance not offered by employer; they use a VA or military hospital; they are covered by other health plans; they haven't needed health insurance; job layoff, loss, unemployment; they are no longer covered by parents; they are not eligible (part time or temporal workers), poor health, illness, age, etc.; some other reason; and too expensive and cannot afford. Models were estimated using all these variables but there is no improvement in results.

[48] More specifically, the changes in the deviation from the mean occupational risk (across all industries) must be uncorrelated with changes in the error term from the regression estimating the changes in wage for the FD model. For the FE model, the time-demeaned deviation from the mean occupational risk (across all industries) must be uncorrelated with time-demeaned error term from the regression estimating the time-demeaned in wage.

FE and IV-FD models based on the second type of instrumental variables. As presented in table 28, *dif_rocc* and *nohi_reason* are significant at 1% and 5%, respectively in the first-stage FE model, and only *dif_rocc* is significant at 1% in the first-stage FD model. Note that most of industry dummy variables are significant risk determinants at the 1% level. In addition, our firm variables, *empsize*, and *hifull*, are also significant risk determinants at the 1% and 10 % levels, respectively. These strong correlations of the risk variable between industry dummy variables and firm-side variables confirm the importance of including these variables in the hedonic wage model, as discussed earlier.

As shown in table 29, the second stage IV-FE model and IV-FD models show the coefficients of the risk variable to be 0.0112 and 0.0077, respectively. The Hausman test for endogeneity results, shown in the *Endogeneity test* row, indicates that these coefficients are not significantly different from those in the FD and the FE models. These results suggest that there is no endogeneity bias in the FD and FE models resulting from contemporaneous correlation. The Sargan statistics, which evaluates the over-identification restriction, fails to reject the null hypothesis. Failing the null hypothesis of the over-identifying restriction indicates that the current set of instruments is valid, although this may be due to the low power of the test (Wooldridge 2001). Nevertheless, the coefficient estimates in the IV-FE and IV-FD models are similar to each other, which indicate that the models are well-specified. There is no significant change among non-risk variables when we estimate the IV-FE and IV-FD models as compared to the FE and FD models.

There may be a question about using *dif_rocc* as an instrument. The instrument should correlate with risk variable and not correlate with error terms in the FD and FE models. As mentioned earlier, the variable *dif_rocc* is not a valid instrument if worker's deviation from the mean risk level within a same occupation is correlated with the error term. Table 30 shows the second-stage results of the IV-FE and IV-FD estimators when we omit *dif_rocc* as an instrument in the first stage. The top half of table 30 presents key results from the second stage of the IV-FE and IV-FD models and the second half shows key results from the first stage. None of the external instruments are significant in either the FD or FE models. The Anderson statistics, which test for the relevance of instruments, fail to reject null hypothesis that indicates that the correlation between the risk variable and external instruments are weak (Baum, 2006). Weak instruments generally make the estimators inconsistent, and increase the standard error of the estimator (Wooldridge 2001). The omission of *dif_rocc* apparently makes the instruments weak and the IV-FD and IV-FE estimators become insignificant. In addition, the IV-FD estimator changes sign indicating that the estimators may indeed be inconsistent. Including additional variables indicating why workers do not have health insurance (see footnote 47), does not improve the results.

It is difficult to say with certainty whether *dif_rocc* is a valid instrument. The conditions be met are quite complicated in this context. Again, the conditions for the FD model are that the changes in the deviation from the mean occupational risk (across all industries) must be uncorrelated with changes in the error term from the regression estimating the changes in wage.[49] There is not an intuitive story as to why this condition might hold. However, there is not a clear argument against its validity either. Furthermore, the closeness of estimated IV-FE and IV-FE estimators when *dif_rocc* is included as an instrument, as well as the failing to reject Sargan statistics in these models,

---

[49] See footnote 48 for the condition of valid instrument for the FE model.

provide some confidence that the models are well-defined. Therefore, we draw our conclusions relying on the results from IV-FE and IV_FD models which include *dif_rocc* as an instrumental variable.

Sensitivity Analysis.

There is a concern that the difference of the OLS and panel estimators may be due to the systematic difference between job changers and non-job changers. In the pooled OLS model, the variation in the risk variable from both job changers and non-job changers contribute to estimate the risk coefficient. On the other hand, since the risk change is zero for all non-job changers, the variation to estimate the risk coefficient in the panel models comes from only the job changer sample. It is possible that job changers and non-job changers may be systematically different in terms of the observable characteristics such as risk taking behavior or age, as well as the unobservable characteristics, such as risk preference or the level of job competence. If job changers and non-job changers are systematically different, then there are two factors contributing the difference between the panel risk estimators and the OLS risk estimator. One is time-invariant worker heterogeneity (within sample heterogeneity), and the other is worker heterogeneity between job changers and non-job changers.

Table 31 shows the summary statistics of key variables of job changers and non-job changers. Job changers are defined as workers who changed jobs at any point in the 1996 Panel. The main noticeable, but relatively minor, differences between job changers and non-job changers are in their average age, union membership and the availability of full employer-provided health insurance. The average age of non-job changer is 39 years old while that of job changers is 36 years old. This difference makes sense because young people tend to change jobs more often. There is a slightly higher proportion of people who are married, but a slightly lower proportion of people who have kids under age 18 among non-job changers as compared to job changers. These differences are likely due to the difference in age between job changers and non-job changers. The non-job changer sample has a higher level of union membership rate and more access to the employer provided full health insurance as compared to job changers. In addition, non-job changers earn slightly higher hourly wage than job changers.

To examine the effect of potential heterogeneity between job changers and non-job changers on the estimating results, we re-estimate the pooled OLS model and panel models under two hypotheses. The first hypothesis is that the job changers and non-job changers face different hedonic wage schedules for occupational risk compensation, but both samples have the same sample distribution in terms of worker/job characteristics other than risk. The second hypothesis is that the job changers and non-job changers may face different hedonic wage schedules for occupational risk compensation and both samples have different sample distributions in terms of other worker/job characteristics.

To examine the first hypothesis, we estimate the following pooled OLS model.

$$y_i = \beta_1 r_i + \beta_2 r_i \times JC + X_i \gamma + \mu_i, \tag{23}$$

where all variables are as described in equation 15 and *JC* is a dummy variable, where 1 indicate a job changer and 0 indicate a non-job changer. This model allows job changers and non-job changers to face different hedonic wage curve for occupational risk, but assumes all other variables have the same sample distribution. The estimation results show that the risk coefficient for job changers is 0.0068 (SE=0.0031) and for non-job changers is 0.0238 (SE=0.0030) and they are significant at 1% and 5% level, respectively. The 95% confidence interval of these two risk coefficients do not overlap. This results strongly indicates that job changers and non-job changers face different hedonic wage schedule for risk. In addition, we do not find a significant difference between pooled OLS risk estimator for job changers and panel risk estimators presented in table 27, which indicates that there is no time-invariant worker heterogeneity bias in the pooled OLS model when based on job changers only.

Now, we assume that the job changers and non-job changers may face a different hedonic wage schedule for risk compensation and also assume that they may have differences in other worker/job characteristics that correlate with the risk variable. To examine this hypothesis, we estimate the pooled OLS and panel models with only the job-changer sample. Table 32 shows the regression results. The risk coefficient in the pooled OLS model is 0.0207 and is significant at the 1% level. The 95% confidence interval of the risk coefficient is between 0.0140 and 0.0275. As shown in table 27, the risk coefficient in the pooled OLS model based on both job-changer and non-job changer sample is 0.0167 with a 95% confidence interval of 0.0118 and 0.0215. When we estimate the pooled OLS model only with non-job changer sample, we obtain the risk coefficient of 0.0130 with a 95% confidence interval of 0.0062 and 0.0197. Thus, including the non-job changers lowers the estimated risk coefficient in the pooled OLS somewhat, but the difference between the coefficient estimates from the three models is not statistically significant at the 5% level. As expected, the risk coefficients from fixed effect and first difference models do not show a significant change by excluding non-job changers as shown in table 32. This result reinforces the original conclusions that the time-invariant worker heterogeneity biases the pooled OLS risk estimator upward.

Since these two hypothesis leads to opposite conclusions, we need to carefully examine which hypothesis is more appropriate. To test the appropriateness of the first hypothesis, we estimate the model which adds interaction terms between job changing status and all variables in the pooled OLS model (such as *risk*×*JC*, *age*×*JC*, *college*×*JC* etc.). We test if the coefficients of all risk interaction terms are jointly different from zero. We obtain the F-statistics of 4.34 and the p-value of 0.0040. The test result suggests that it would be reasonable to assume that job changers and non-job changers are systematically different in terms of worker/job characteristics. Therefore, we rely estimation results from the second hypothesis for our conclusions.

With only the job changer sample, we estimate the VSL of $7.78 million, $2.59 million, and $1.70 million from the pooled OLS, FE model and FD model, respectively. This result indicates that the use of the OLS model biases the risk coefficient upward significantly due to time-invariant worker heterogeneity. Table 33 shows the 2SLS panel models with the second type of instruments with only the job changer sample. We do not find endogneiy problem in our revised panel models.

Next, we examine the potential bias associated with the market disequilibrium. If the workers who tend to change jobs are ones who are out of equilibrium, and they change jobs so that they move towards an equilibrium position, then panel models do not provide consistent estimators. If this is the case, the panel estimators represent the movement of workers from disequilibrium to the equilibrium, and not the static hedonic wage schedule. Herzog and Schlottman (1990) use industry switching models to examine this disequilibrium hypothesis. In this study, we compare hedonic wage schedule of before and after job change for the job changer sample.

Table 34 shows the summary of the sensitive analysis results. The last row and the third from the last row shows the OLS results of the before job change sample and after job change sample, respectively. The risk coefficient of before job change sample and after job change sample is 0.0164 and 0.0234, respectively. The difference in the risk coefficients is not statistically significant at the 5% level. Thus, similarity in the two estimates suggests that workers are not moving towards different HW schedules by changing jobs. Of course this is not a concrete test for market disequilibrium since the OLS estimation results may be biased for other reasons as shown in this chapter. However, unless there is a strong reason why we should believe there is a different direction or degree of bias in the before job change sample and the after job change sample, this simple test provides useful information about the market disequilibrium hypothesis for our sample.

Conclusions.

This study aims to identify the endogeneity bias in previous cross-section HW studies by combining the panel data approach and the IV approach. The endogenous nature of risk variable in HW models has been a major concern in using the wage/risk premium estimators in policy analysis (Viscusi & Aldy 2003). This study uses two stage least square (2SLS) panel models to control for endogeneity bias resulting from omitted time-variant and time-invariant individual heterogeneity, measurement error associated with the risk variables and simultaneity between the wage and risk variable.

We collect a sample of workers from a large national panel study, the 1996 SIPP. Our occupational fatal risk rates vary by occupation and industry. This fatal risk measure provides important variation of the level of risk when workers change their occupation, and enable us to estimate the wage/risk premium with panel models.

Our analyses find that there is a significant upward bias due to the omitted time-invariant worker heterogeneity. The estimated VSL after controlling for the omitted time-invariant worker heterogeneity is between $1.7 million-$2.5 million, which is nearly a half of the VSL estimate when we do not control for the omitted time-invariant worker heterogeneity. Our analysis of two stage least square panel models show that there is no endogeneity bias in our panel models resulting from time-variant worker heterogeneity measurement error or simultaneity between risk variable and wage.

Our finding is similar to the finding of Kniesner et al. (2005), which also use panel models to estimate wage/risk premium. They found that the estimated VSL after controlling for the omitted time-invariant worker heterogeneity is $6.7 million, which is about a third of their VSL estimate without controlling for the omitted time-invariant worker heterogeneity.

The difference in the absolute value of the VSL between the Kniesner et al. study and our study comes from the different model specifications and the different wage level of the sample. We find that industry dummy variables and firm-side variables that are omitted from Kniesner et al. (2005) are significant wage determinants, and also correlate with the risk variable. We show that omitting these variables bias the risk coefficient upward by about 50% in both cross-section and panel models. Interestingly in footnote 9 in page 14, Kniesner et al. (2005) report that the VSL estimate becomes $4.4 million when they include the one-digit level industry dummy, which is also about 50% smaller than their estimate of the VSL when excluding the one-digit level industry dummy.

However, in their text, they express concern about multicolinearity between the risk and industry dummy variables and do not consider the model with industry dummies further.

Our sample of workers has a lower than average wage, while Kniesner et al.'s (2005) sample has a higher than average wage. This divergence of the average wage level may contribute to the disparity between our VSL estimate and their VSL estimate. To further explore this issue, we re-estimate the FE and IV-FE models with an adjusted SIPP sample. We first exclude workers who are earning less than $9.7 per hour so that the sample average wage matches the U.S. average wage, $16 (see chapter 3). This leaves 111,723 observations in the sample. The estimated FE shows the coefficient of risk variable 0.0062, and a VSL of $2.0 million. We further reduce our sample of workers to those earning more than $16 per hour so that our sample average wage matches the average wage level in the Kniesner et al. sample. This leaves 44,605 observations in the sample. However, we do not obtain significant risk estimator with this sample in the FE model. Although these results are only suggestive since we only match the mean, not the variance in wage across samples, they do suggest that perhaps the industry controls play a more important role in the difference our estimates as compared to the wage differences.

It is important to remember that our sample is not a representative sample of workers in the U.S. (see chapter 3), and we must be cautious to apply our VSL estimation to policy analysis. However, it is not likely that the uniqueness of our sample would undermine our conclusions about the existence of endogeneity bias in the cross-section HW models and the critical importance of including firm-side components that Kniesner et al. fail to explore.

As a sensitivity analysis, we consider whether our comparison of pooled OLS and panel models are appropriate to examine the bias due to time-invariant worker heterogeneity in pooled OLS model. If there is heterogeneity between the job-changer and non-job changer sample, the difference between pooled OLS and panel model estimators comes from two sources; the worker heterogeneity and heterogeneity between job-changers and non-job changers. When we estimate models only for job-changers, we find larger difference between pooled OLS and panel estimators. This result reinforces our original observation of significant upward bias in pooled OLS estimators due to the time-invariant worker heterogeneity.

We also examine our underlying assumption that workers are in equilibrium in any period of time. In our sensitive analysis, we do not find a significant HW schedule change between before and after job change for the job changer sample. We conclude that workers do not move between different HW schedules, but move along the same HW schedule when they change jobs, and thus the our underlying assumption is valid.

Table 27

Cross-section, Fixed Effect and First Difference Regression Results

	OLS (clustered)	(SE)	Fixed Effect	(SE)	First Difference	(SE)
risk	0.0167***	0.0024	0.0094***	0.0015	0.0062***	0.0022
age	0.0303***	0.0010	0.0556***	0.0020	0.0156***	0.0027
age2	-0.0003***	0.00001	-0.0007***	0.00001	-0.0001***	0.00003
ugdeg	0.2022***	0.0089	0.1097***	0.0179	0.0183	0.0259
college	0.1373***	0.0058	0.0199***	0.0137	0.0026	0.0195
hsgrad	0.0785***	0.0053	0.0150***	0.0120	0.0009	0.0158
hispanic	-0.0095***	0.0059				
blacknh	-0.0569***	0.0055				
female	-0.1301***	0.0043				
workov	0.0637***	0.0037	0.0147***	0.0015	0.0071***	0.0012
union	0.1983***	0.0049	0.0515***	0.0023	0.0163***	0.0019
kids18	0.0058***	0.0017	0.0014	0.0013	0.00008	0.0017
married	0.0739***	0.0039	0.0171***	0.0031	0.0056	0.0038
hipart	0.1426***	0.0036	0.0350***	0.0016	0.0116***	0.0014
hifull	0.1513***	0.0043	0.0383***	0.0018	0.0118***	0.0015
empall	0.0472***	0.0033	0.0131***	0.0014	0.0060***	0.0014
empsize	-0.0472***	0.0035	-0.0149***	0.0016	-0.0070***	0.0015
neast	-0.0187***	0.0063	0.0174	0.0195	0.0685	0.0272
midwest	-0.0421***	0.0055	-0.0095	0.0155	0.0623	0.0223
south	-0.0962***	0.0053	-0.0501***	0.0144	0.0077	0.0210
urban	0.0709***	0.0043	0.0111***	0.0037	0.0026	0.0048
agind	-0.1309***	0.0166	-0.0418***	0.0111	-0.0241***	0.0181
constind	0.0112	0.0114	0.0175*	0.0084	-0.0124	0.0141
tcuind	-0.0399***	0.0107	-0.0122	0.0085	-0.0590***	0.0145
trdind	-0.2205***	0.0089	-0.0957***	0.0072	-0.1135***	0.0125
servind	-0.0933***	0.0086	-0.0485***	0.0069	-0.0714***	0.0121
manufind	-0.0955***	0.0091	-0.0105**	0.0074	-0.0432***	0.0127
craftocc	0.2605***	0.0154	0.0717***	0.0088	0.0611***	0.0121
profocc	0.3151***	0.0160	0.0868***	0.0091	0.0733***	0.0124
techocc	0.1849***	0.0152	0.0558***	0.0087	0.0501***	0.0120
servocc	0.0075	0.0156	-0.0227***	0.0089	-0.0215*	0.0123
laborocc	0.1130***	0.0151	0.0343***	0.0086	0.0192*	0.0116
Constant	1.5049***	0.0251	1.3747***	0.0630	0.0103***	0.0005
N (# group)	166,362	(34,847)	166,362	(34,847)	113,343	(26,269)
R2 (overall)	0.45		0.06		0.26	
VSL (million$)	4.60		2.59		1.71	

***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

Table 28

First Stage Regression Results With Second Set of Instrumental Variables.

	Fixed Effect model	(SE)	First Difference model	(SE)
age	0.0075***	0.0024	0.0005	0.0023
age2	-0.00007***	0.00002	1.47×10^{-6}	0.00002
ugdeg	0.0455**	0.0206	0.0085	0.0220
college	0.0214	0.0158	-0.0070	0.0165
hsgrad	0.0122	0.0138	-0.0129	0.0134
workov	0.0048***	0.0017	0.0039***	0.0010
union	0.0062**	0.0026	0.0024	0.0016
kids18	-0.0011	0.0015	-0.0006	0.0014
married	0.0050	0.0035	0.0049	0.0032
hipart	-0.0004	0.0018	0.0008	0.0011
hifull	-0.0039*	0.0021	0.0003	0.0013
empall	-0.0025	0.0017	-0.0026**	0.0012
empsize	0.0149***	0.0018	0.0038***	0.0013
neast	0.0254	0.0225	-0.0530	0.0231
midwest	-0.0110	0.0179	0.0321*	0.0189
south	-0.0053	0.0165	-0.0010	0.0179
urban	0.0037	0.0042	-0.0059**	0.0040
agind	0.4900***	0.0126	0.3973***	0.0153
constind	0.8854***	0.0094	0.8795***	0.0117
tcuind	0.3793***	0.0097	0.3107***	0.0123
trdind	0.0008	0.0083	-0.0008	0.0106
servind	-0.0512***	0.0080	-0.0630***	0.0103
manufind	-0.2171***	0.0085	-0.2015***	0.0107
craftocc	-0.6820***	0.0101	-0.8254***	0.0102
profocc	-1.3007***	0.0105	-1.4503***	0.0105
techocc	-1.3397***	0.0102	-1.5134***	0.0101
servocc	-1.0315***	0.0103	-1.1550***	0.0104
laborocc	-0.4030***	0.0100	-0.5644***	0.0099
Dif_rocc	0.6529***	0.0015	0.6790***	0.0016
Inc_other	2.42×10^{-7}	1.06×10^{-7}	-2.98×10^{-7}	6.14×10^{-7}
Inv_all	0.00001	9.43×10^{-6}	1.77×10^{-6}	5.70×10^{-6}
N_SS	0.0034	0.0025	-0.0019	0.0017
Nohi_reason	0.0281*	0.0171	0.0060	0.0091
Constant	1.3984***	0.0726	0.0004	0.0004
R2 (overall)	0.78		0.72	

***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

Table 29

Second Stage (IV-FE and IV-FD) Regression Results: Second Set of Instruments.

	IV-Fixed Effect	(SE)	IV-First Difference	(SE)
risk-hat	0.0112***	0.0021	0.0077***	0.0028
age	0.05567***	0.0020	0.0156***	0.0027
age2	-0.0007***	0.00001	-0.0001***	0.00003
ugdeg	0.1096***	0.0179	0.0183	0.0259
college	0.0198	0.0137	0.0026	0.0195
hsgrad	0.0149	0.0120	0.0009	0.0158
workov	0.0147***	0.0015	0.0071***	0.0012
union	0.0515***	0.0023	0.0163***	0.0019
kids18	0.0014	0.0013	0.00008	0.0017
married	0.0171***	0.0031	0.0056	0.0038
hipart	0.0350***	0.0016	0.0110***	0.0013
hifull	0.0383***	0.0018	0.0118***	0.0015
empall	0.0131***	0.0014	0.0060***	0.0014
empsize	-0.0149***	0.0016	-0.0070***	0.0015
neast	0.0173	0.0195	0.0686	0.0272
midwest	-0.0096	0.0155	0.0624*	0.0223
south	-0.0502***	0.0144	0.0078	0.0210
urban	0.0111***	0.0037	0.0026	0.0048
agind	-0.0418***	0.0111	-0.0257	0.0182
constind	0.0150*	0.0085	-0.0146	0.0143
tcuind	-0.0137	0.0085	-0.0601***	0.0146
trdind	-0.0956***	0.0072	-0.1134***	0.0125
servind	-0.0483***	0.0069	-0.0711***	0.0121
manufind	-0.0099	0.0074	-0.0427***	0.0127
craftocc	0.0722***	0.0088	0.0621***	0.0121
profocc	0.0878***	0.0091	0.0746***	0.0125
techocc	0.0568***	0.0088	0.0515***	0.0121
servocc	-0.0219**	0.0089	-0.0203*	0.0123
laborocc	0.0339***	0.0086	0.0195*	0.0116
Anderson LR statistics	P<0.01		P<0.01	
Sargan statistics	P=0.83		P=0.94	
Endogeneity test	P=0.19		P=0.786	
N (# group)	157,784	(26,269)	113,343	(24,142)
R2 (overall)	0.08		0.01	

***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

Table 30

IV-FE and IV-FD Results With First Set of Instruments.

	IV-Fixed Effect	(SE)	IV-First Difference	(SE)
Selected second-stage results				
risk-hat	0.0077	0.3052	-0.1210	0.6047
Anderson LR statistics	P<0.46		P<0.81	
Sargan statistics	P=0.68		P=0.86	
Endogeneity test	P=0.99		P=0.83	
N (# group)	157,784	(26,269)	113,343	(24,142)
R2 (overall)	0.08		0.01	
Selected first-stage results				
inc_other	1.01×10^{-6}	1.61×10^{-6}	-2.91×10^{-7}	9.76×10^{-7}
inv_all	0.00001	9.43×10^{-6}	8.76×10^{-7}	9.06×10^{-6}
N_SS	0.00166	0.00380	-0.0030	0.0028
nohi_reason	0.03748	0.0258	-0.0048	0.0145
R2 (overall)	0.53		0.30	

Table 31

Summary Statistics of Key Characteristics of Job Changers and Non-job Changer Sample.

	Non-job changers		Job changers	
	Mean	SD	Mean	SD
wage	13.8676	6.2192	12.5094	5.3895
risk	0.5643	0.9690	0.5508	0.9410
age	39.4184	11.58	36.3091	11.0654
college	0.3283	0.4696	0.3363	0.4724
hsgrad	0.4134	0.4924	0.4278	0.4947
hispanic	0.1281	0.3342	0.1413	0.3483
blacknh	0.1332	0.3398	0.1419	0.3489
othrace	0.0385	0.1924	0.0436	0.2042
female	0.4628	0.4986	0.4517	0.4976
workov	0.1909	0.3930	0.1892	0.3916
union	0.2252	0.4177	0.1580	0.3647
kids18	0.7627	1.1019	0.8262	1.1086
married	0.5820	0.4932	0.5361	0.4986
hipart	0.4542	0.4979	0.4346	0.4957
hifull	0.2216	0.4153	0.1806	0.3847
empall	0.5676	0.4954	0.5581	0.4966
empsize	0.2948	0.4559	0.2990	0.4587

Table 32

Cross-section, Fixed Effect and First Difference Regression Results With Only Job Changer Sample.

	OLS (clustered)	(SE)	Fixed Effect	(SE)	First Difference	(SE)
risk	0.0207***	0.0034	0.0090***	0.0017	0.0059**	0.0023
age	0.0303***	0.0015	0.0629***	0.0033	0.0158***	0.0044
age2	-0.0003***	0.00002	-0.0008***	0.00002	-0.0002***	0.00005
ugdeg	0.1790***	0.0135	0.0781***	0.0242	-0.0107	0.0388
college	0.1310***	0.0083	-0.0153	0.0187	-0.0243	0.0312
hsgrad	0.0801***	0.0076	-0.0082	0.0164	-0.0156	0.0253
hispanic	-0.0928***	0.0084				
blacknh	-0.0453***	0.0079				
othrace	-0.0378***	0.0146				
female	-0.1316***	0.0061				
workov	0.0638***	0.0053	0.0208***	0.0024	0.0097***	0.0018
union	0.1907***	0.0075	0.0779***	0.0036	0.0264***	0.0031
kids18	0.0026	0.0024	0.0010	0.0020	-0.0020	0.0025
married	0.0791***	0.0056	0.0288***	0.0046	0.0140**	0.0060
hipart	0.1390***	0.0050	0.0477***	0.0024	0.0159***	0.0021
hifull	0.1469***	0.0061	0.0516***	0.0029	0.0141***	0.0024
empall	0.0463***	0.0046	0.0170***	0.0021	0.0087***	0.0021
empsize	-0.0422***	0.0049	-0.0216***	0.0023	-0.0124***	0.0022
neast	-0.0243***	0.0093	0.0142	0.0252	0.0587	0.0398
midwest	-0.0381***	0.0081	0.0021	0.0204	0.0894***	0.0325
south	-0.0949***	0.0077	-0.0456**	0.0188	0.0192	0.0301
urban	0.0743***	0.0062	0.0130**	0.0052	0.0027	0.0067
agind	-0.1149***	0.0249	-0.0329***	0.0125	-0.0227	0.0195
constind	-0.0063	0.0168	0.0245***	0.0095	-0.0099	0.0152
tcuind	-0.0567***	0.0167	-0.0095	0.0096	-0.0588***	0.0156
trdind	-0.1963***	0.0139	-0.0874***	0.0082	-0.1118***	0.0135
servind	-0.0873***	0.0135	-0.0409***	0.0079	-0.0695***	0.0131
manufind	-0.0810***	0.0142	-0.0079	0.0084	-0.0439***	0.0136
craftocc	0.2574***	0.0225	0.0669***	0.0099	0.0596***	0.0130
profocc	0.3020***	0.0233	0.0814***	0.0103	0.0717***	0.0133
techocc	0.1869***	0.0223	0.0513***	0.0099	0.0490***	0.0129
servocc	0.0211	0.0228	-0.0261***	0.0101	-0.0228*	0.0132
laborocc	0.1113***	0.0221	0.0301***	0.0097	0.0180	0.0125
Constant	1.4790***	0.0365	1.2468***	0.0971	0.0126***	0.0008
N (# group)	72,658	(11,294)	72658	(11294)	49102	(11294)
R2 (overall)						
VSL (million$)	7.78		2.59		1.70	

***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

Table 33

IV-FE and IV-FD Results With Second Type of Instruments: Job-changer Sample Only.

	IV-Fixed Effect	(SE)	IV-First Difference	(SE)
Selected second-stage results				
risk-hat	0.0108***	0.024	0.0074**	0.0030
Anderson LR statistics	P<0.01		P<0.01	
Sargan statistics	P=0.69		P=0.73	
Endogeneity test	P=0.24		P=0.66	
N (# group)	72,658	(11,294)	4,9102	(10,020)
R2 (overall)	0.11		0.01	
Selected first-stage results				
dif_rocc	0.6528***	0.0023	0.6789***	0.0024
inc_other	4.39×10^{-7}	2.29×10^{-6}	-7.11×10^{-7}	1.48×10^{-6}
inv_all	0.00004	9.43×10^{-6}	2.16×10^{-6}	0.00002
N_SS	0.0065	0.0050	-0.0038	0.0038
nohi_reason	0.0523	0.0349	0.0127	0.0200
R2 (overall)	0.78		0.72	

Table 34

Summary of Estimated Coefficients.

	Coefficient (standard error)	95% confidence interval	
Job changers-FD	0.0059 (0.0023)	0.00123	0.0105
All-FD	0.0062 (0.0022)	0.0018	0.0105
Job changers-FE	0.0090 (0.0017)	0.0055	0.0125
All-FE	0.0094 (0.0015)	0.0064	0.0124
Non-job changers OLS	0.0130 (0.0034)	0.0062	0.0197
All-OLS	0.0167 (0.0024)	0.0119	0.0215
Job changers OLS - before	0.0164 (0.0035)	0.0096	0.0233
Job changers OLS (before+after)	0.0207 (0.0034)	0.0140	0.0275
Job changer OLS-after	0.0234 (0.0038)	0.0159	0.0309

Chapter V

Transferability of VSL: Hedonic Wage Analysis

This study examines the transferability of the Value of Statistical Life (VSL) between different policy contexts using labor market data. This study is motivated by the persistent concern regarding the use of VSL estimates based on fatal risks which are not directly related to the policy objective. For example, U.S.EPA relies on the labor market studies to obtain the VSL to evaluate air pollution control program or fatal cancer reduction programs. This exercise faces several criticisms; one of them comes from the concern that individual's may have different preferences towards heterogeneous risks (USEPA, 1997, 2005; S. A. B. USEPA, 2000). If individuals are sensitive to the qualitative characteristics of risk, such as dread, fear, and controllability, they may prefer certain risks over others even though both risks have a same probability of death. If this is the case, individuals will place different values on risk reductions of different types. This implies that individuals may place different values on reducing environment-oriented risks as compared to occupational risks, and the VSL obtained in the context of occupational risk is not a valid VSL to apply in the environmental policy evaluation.

Early psychology studies did find that individuals perceive fatal risks differently depending on the perceived risk characteristics. For example, Slovic et al. (1980) find that lay people judge the degree of "risk" of death not only from the probability of death but also the characteristics of death. Slovic et al. (1980) survey different groups and ask them to order the risk of 30 different activities, including nuclear power, motor vehicle, smoking, vaccinations etc. They find that the order made by a group of experts matched the actual frequency or probability of events, while the order of lay people does not. It was possible that the perceived probability of lay people itself was biased by different reasons, such as an exposure to media coverage (Combs & Slovic, 1979). However, Slovic et al. also find that the order that lay people assigned to the risk of death from various activities did not match with their perceived probability of each event either. Slovic et al. further examine what factors drive the disparity between lay people's risk perception and their perceived probability of event. Key characteristics that seem to affect individuals' risk perception include the degree of control over the risk, the amount of dread involved, or potential to threat the future generations (Slovic et al., 1980).

Several contingent valuation (CV) studies further examine the relationship between risk characteristics and individual's willingness to pay (WTP) to reduce fatal risks. There are three main approaches in these studies. The first approach elicits individual's WTP to reduce fatal risks that arise from different contexts, such as air pollution or traffic accidents (Cookson, 2000; McDaniels et al., 1992; Savage, 1993). The second approach elicits individual's preference towards different policy programs which reduce different types of fatal risk but the cost effectiveness of the policy is identical (Cookson, 2000; Cropper & Subramanian, 1999). The third approach elicits the individual perception on the equivalent number of lives saved in different policy contexts to a certain policy program (Chilton et al., 2002; Cookson, 2000; Jones-Lee & Loomes, 1995). These studies also ask individuals to rate the degree of qualitative characteristics of each risk according to their perception. The elicited WTPs or preference towards certain policy are statistically or non-statistically related with the perceived risk characteristics, and estimate the marginal effect of each risk characteristic on the WTP.

Jones-Lee and Loomes (1995), Cropper and Sabramanian (1999), and Cookson (2000) found that individuals tend to exhibit a lower WTP (or place lower priority) to reduce highly controllable and voluntary risks (e.g., automobile accident risks) than to reduce other types of risks (e.g., air pollution). Also, McDaniel et al.(1992) and Savage (1993) found that individuals have a higher WTP for reducing risks that involve a high degree of dread as compared to other types of risk.

While there are a number of survey-based studies suggesting that individuals value a reduction of different types of risks significantly differently, we have little evidence from revealed preference studies in these regards. There are two major advantages of using the revealed preference methods. One advantage is that this method bases on the data from individuals' actual behavior and eliminates the hypothetical bias which often pertains to survey studies (Cummings & Taylor, 1999). The second advantage is that the labor market approach may mitigate the problem associated with the subject's inability to understand small risk levels, which again is often an issue with survey methods (Corso, Hammitt, & Graham, 2001; Hammitt & Graham, 1999). Benjamin (2001) shows that on average, individuals can estimate their *personal* probability of death from various sources quite accurately. Since occupational risk is a personal risk for workers, it is reasonable to assume that *on average*, these workers understand their probability of death at work reasonably well.

This study attempts to provide insights into whether or not heterogeneous workplace risks play an important role in workers decision-making using labor market data. In particular, we will focus on the risk/wage tradeoff between two very different risks: violent assault (homicide) risks and non-violent risks. Although the focus is on the individual's behavior within the labor market, this study is potentially important to evaluate the transferability of VSL across different policy contexts. If individuals exhibit different WTP to avoid different types of risks within the same policy context such as workplace safety, it is hard to justify the application of VSL transfer between different policy contexts.

The next section describes previous work which this study improves upon, followed by a presentation of the empirical model. Finally, results and conclusions will be presented.

Literature Review

To date, there is only one study that attempts to estimate the wage/risk premium for different types of risk using the hedonic wage model. Scotton and Taylor (2006) construct an occupational fatality rate which varies by the cause of death from the public use sample of the Census of Fatal Occupational Injuries (CFOI). Using this data, they estimate wage/risk premiums for different types of risk faced by a broad sample of workers.

The results of their hedonic wage model are puzzling. They find implausibly large risk premiums or large *negative* premiums for violent assault risks. The authors suggest that the reasons they may have failed to estimate a theoretically consistent wage premium are: 1) the objective measures of risk they used may be different from the workers' subjective measure of risk where the type of death is a rare event, and 2) the hedonic model they estimated may have failed to take into account the unobserved non-risk characteristics of the job.

Also, Scotton and Taylor did not take into account the spatial heterogeneity of occupational risk. For instance, workers in large cities tend to face higher crime and traffic accident rates, and thus face higher occupational fatal risks than those in rural areas. The practice of assigning the same level of risk to workers in the same occupation regardless of location may have biased their result.

In this study, we will improve the study design of Scotton and Taylor by using a different sample of workers as well as creating new, location-specific different risk data to estimate the wage/risk premia for two different types of risks: violent-assault risk and non-violent risk.

To avoid measurement error caused by the disparity between an objective and subjective measure of risk, we will use a sample of workers who face either high violent assault or high non-violent (or both) risks routinely as part of their job. Benjamin (2001) shows that individuals perform better in estimating their personal risk level when the actual risk level faced is high. It is hoped that individuals who face high level of fatal risk routinely as part of their job are likely to understand the objective risk level of their job correctly. Also, to minimize the bias caused by the unobserved job characteristics, we will use a sample of workers in a homogeneous occupation which requires very similar job duties.

To fulfill these requirements, we will use a sample of occupational drivers, which includes taxi, truck, sales and bus drivers. Occupational drivers face higher traffic-accident risks on the job than other occupations, yet the type of risks faced varies across driver-types. As discussed in chapter 3, the main causes of driver death are violent assaults and traffic-related accident. Truck drivers have the highest traffic-related fatalities among driving occupations. Taxi drivers, on the other hand, have the highest homicide rate among those in driving occupations. Indeed, the risk of death from violent assault for taxi drivers is the highest among all occupations, nearly four times more than the homicide rate of police officers (Knestaut, 1997). In addition, the fatal risk rates from different causes varied significantly by the geographic area in which drivers reside.

Slovic et al. (1980) show that lay people considers *crime* as a highly uncontrollable, involuntary and a highly dreaded risk. In addition, lay people generally consider the fatal risk involving motor vehicles as relatively controllable, voluntary and with less dread. Therefore, the comparison between violent-assault and non-violent risks for occupational drivers may be an excellent case in which we can examine the degree to which risks with different characteristics, especially in terms of dread and control involved, are valued differently by individuals.

Empirical Model.

We will use a cross-section hedonic wage model to estimate the risk/wage premium for a specialized sample of workers for each cause of fatal risk. The estimating hedonic wage model is:

$$\ln wage_i = a_0 + a_1 vrisk_{oj} + a_2 nvrisk_{oj} + a_3 vinj_{oj} + a_4 nvinj_{oj} + X_i\beta + W_o\gamma + Z_j\delta + \varepsilon_i \tag{24}$$

where i denotes an individual worker in occupation o in area j. The variables for occupational risk are as follows; *vrisk* is the fatal risk from a violent assault, *nvrisk* is fatal risk from a non-violent event, *vinj* is the risk of non-fatal injury (injury risk) from violent interactions and *nvinj* is the risk of non-fatal injury from non-violent causes. These risk rates vary by occupation (o) and geographic area (j). X is a vector of relevant individual characteristics, W is a vector of relevant job characteristics other than occupational risk, Z is regional characteristics which affect wage levels in each area, and ε is an error term.

From the estimated HW model, we test if the wage/risk premiums for different types of fatal risk (a_1 and a_2) and for different types of injury risk (a_3 and a_4) are statistically different from each other. The wage/risk premium for violent assault fatal/injury risk are expected to be significantly larger than that for non-violent fatal/injury assault risk due to the nature of violent risk with higher level of perceived uncontrollability and dread. We assume the same wage-risk compensation schedule for each occupation.

The choice of which geographic level to create the risk rates to be used may be important to estimate HW models accurately. If the drivers work only within the MSA, the MSA level risk rates are the relevant risk measure. If, however, the drivers work all over the state in which they reside, then a state-level risk rate is a more relevant risk measure. It would be reasonable to assume that each driver works within and around the MSA in which they reside. In this case, drivers would form their perception about their risks by weighting the MSA level risk and the state level risk according to the time they spend inside and outside the MSA.

However, there is no record available from the CPS on how each driver in the sample allocates his/her time inside and outside of the MSA. Thus, the choice becomes either using the MSA-level or state-level risks. We examine the sensitivity of results to the use of different geographic level risks in the next section.

Results.

This section presents the results of our analyses. The base results, which are based on the MSA-level violent risk and the state-level non-violent risks created from the 1992-2002 and 1998-2002 risk data, are presented first. In general, the violent fatal risk is higher at the MSA-level than the sate-level, and the non-violent fatal risk is higher at the state level. We include the MSA-level violent fatal risk and the state-level non-violent fatal risk in the base model to avoid the possible underestimation of either risk level drivers face. The summary of risks by type of risk and geographic level are presented in chapter 3.

We first test the sensitivity of the base model to the inclusion of injury risk variables. Additional sensitivity analyses are as follows. The second and third analyses examine the sensitivity of results to the use of risk rates created from the different geographic levels as well as the inclusion of different geographic level dummy variables. The fourth analyses examine the effect of adjusting the standard error in the model to allow for correlation among workers in the same geographic area. We include geographic-level correlation that is different from the geographic level included as dummy variables. For example, if a model includes MSA dummy variables, then correlation is allowed at the state level. Allowing the correlation among observations, so called *cluster*, affects the standard errors of estimates (Sribney 2005).[50] Additional sensitivity analyses examine the effect of excluding the workers who face zero or high

---

[50] The formula to estimate this variance is: $\text{var} = (X'X)^{-1}\left[\sum_{j=1}^{n_c} u_j{}' u_j\right](X'X)^{-1}$

where $u_j = \sum_{j_{cluster}} e_i x_i$ and n_c is the total number of cluster (William Sribney 2005 retrieved March 25, 2007 from http://www.stata.com/support/faqs/stat/cluster.html).

violent fatal risks and examining the effect of including MSA-specific descriptive variables.

Table 35 shows the regression results with the base model and base sample of workers. The fatal risk data is created based on the 1992-2002 CFOI in model 1 and based on 1998-2002 CFOI in model 2. The base sample is occupational drivers who are not self employed, who work full time, who earn more than minimum wage, who work only one job and who are between age 18 and 66. MSAs with less than 100 employees are omitted from the sample.[51] Also, only workers in the continental U.S. are included. The base models (model 1 and model 2) include violent fatal risk, non-violent fatal risk, demographics, occupation and MSA dummy variables. Since MSA-level violent risk is only available for drivers in MSA, we exclude drivers who live outside of MSA in the base model.

All demographic variables show the expected sign. Older workers earn more than younger workers, but at a decreasing rate. College graduates and high school graduates earn more than non-high school graduates, white workers earn more than other workers, U.S. citizens earn more than non-US citizens, male workers earn more than female workers, unionized worker earn more than non-unionized workers, and married workers earn more than single workers. Among occupations, truck drivers earn the highest level of wage, and taxi drivers earn the lowest level of wage.

In both model 1 and model 2, the violent assault fatal risk has a significantly positive coefficient at the 1% level.[52] The non-violent fatal risk has a negative, but not

[51] Including MSAs with less than 100 employments does not change the results.

[52] We also include the squared violent and non-violent risk terms in the model. However, the squared risk terms are not significant in any model estimate, which indicates that the linear risk model is the preferred model.

statistically significant coefficient in model 1 and negative and significant coefficient in model 2. The coefficient of violent assault fatal risk is statistically larger than that of non-violent fatal risk at the 5% level in model 1 and at the 1% level in model 2 according to Wald tests. A significantly larger coefficient for violent fatal risk than for non-violent fatal risk indicates that the workers require higher wage compensation to accept a marginal increase in violent assault fatal risk than to accept a marginal increase of non-violent fatal risk. The estimated VSL based on the violent fatal risk and mean wages in the sample is \$1.8 million and \$5.1 million for model 1 and mode 2, respectively.[53]

We also estimate the VSL based on the *total* risk, which is the sum of violent and non-violent fatal risks. The coefficient of *total* risk is 0.0034 in model 1 and 0.0052 in model 2. They are significant at the 5% level and the 10% level, respectively. The estimated VSL based on total risk and average annual wage is \$1.2 million and \$1.8 million for model 1 and model 2, respectively. The VSL estimates based on *total* risk is compatible with previous VSL estimates and VSL estimated from panel models in chapter 4. Our VSL estimates based on driver sample is lower bound of VSL estimate range (Kochi et al., 2006; Mrozek & Taylor, 2002; Viscusi & Aldy, 2003).

Table 35 shows that the use of different period of risk data makes significant impact on the risk premium estimation. Model 2 has almost three times larger coefficient for violent risk as compared to model 1. As discussed in chapter 3, the risk data created from 1992-2002 CFOI likely contains significant measurement error, which would bias the estimated risk coefficients downward. In the following sections, we discuss the

[53] VSL is calculated as follows: VSL=coefficient of risk variable ×weekly wage ×52 (weeks) ×10,000, where 10,000 is the unit of fatal risk.

sensitivity of model 2 to different model specifications and different geographic level fatal risks.

Table 36 shows the sensitivity of the base result (model 2) to the different model specifications. Model 3 adds the injury risk variables to model 2. Adding injury risk variables slightly reduce the magnitude of coefficient for violent risk and remove the significance of negative coefficient of non-violent fatal risk. Injury risk is significant and it is more theoretically sound to include it in the wage model, indicating that the model 3 is preferred specification to model 2. Thus, the remaining analyses include injury risk variables. In model 3, the difference of coefficients between two fatal risk variables is significant at the 10% level. The violent injury risk has a significantly positive coefficient at the 1% level, and the non-violent injury risk shows a positive but not significant coefficient. The Wald test shows that the coefficient of violent injury risk is statistically larger than that of non-violent injury risk at the 1% level. The estimated value of statistical injury for total injury risk is about $70,000.[54] This is in the line of previous estimated value of statistical injury, which ranges from $30,000 to 360,000 (in 2005 dollars) (Viscusi & Aldy 2003). The value of statistical injury for violent injury risk in this study is $1.0 million.

Model 4 repeats model 3, but replaces the MSA dummy variables with state dummy variables. Using state level dummy variable has little impact on the risk coefficient estimates, but does change comparison across risk types. The Wald test now fails to reject the null of no significant difference between the two fatal risk coefficients

---

[54] The coefficient of total injury risk for model 3 is 1.941 and significant at 1% level. The value of statistical injury is calculated by multiplying 1.941 by annual wage.

at the 10% level. However the difference is significant at 10.5% level. The difference in coefficients of injury variables is still highly significant.

Table 37 shows the sensitivity of estimation results to using different combinations of geographic level that defining the risk variables. Model 5 and model 6 include the MSA-level violent and non-violent risk variables, and model 7 and model 8 include state-level violent and non-violent risk variables. Model 5 and model 7 include MSA dummy variables and model 6 and model 8 include state dummy variables. Note, in Model 5 and model 8, the risk variables vary by the geographic level that is the same level as the included geographic dummy variables. For example, in model 5, risk only varies by MSA and occupation. In this model, including MSA dummy variables reduce the variation of risk variables significantly, since the only variation to estimate the risk/wage premium comes from the variation over the occupation within the same MSA. Thus, it may be more reasonable to include geographic dummy variables which are not the same level in which the risk variables are created. Model 6 and model 7 include the geographic dummy variables that are not same level in which risk variable is created. These models (models 6 and 7) show that two risk coefficients are statistically different from each other at the 5% level. On the other hand, in model 5 and 8, the differences between the two risk coefficients are not significant.

Although the same issue of multicolinearity applies to injury risk variables, which are only created at the state level (and varies by occupation), the results are robust to include state dummy variables. For all models in table 37, the coefficients of violent injury risk are positive and significant at 1% level and the coefficients of non-violent injury are positive and not significant. The Wald tests show that the differences in coefficients of injury variables are significantly different from each other for all models at the 1% level.

To avoid multicolinearity between risk variables and geographic dummy variables, we only focus on the model specifications that have a combination of fatal risk rates and geographic dummy variables that are created at different geographic level, such as model 6 and model 7. Although in model 6, violent injury risks and geographic dummy variables are created at same geographic levels (state level), it seems there is no significant bias by doing so. Therefore model 6 remains in our set of preferred models.

Table 38 repeats the specifications of model 6 and model 7, but we allow for correlation among workers in a same geographic area.[55] This allows us to examine the sensitivity of results to allowing the correlation within a geographic unit, but assuming there is no correlation across geographic units. For model 9, the model includes state dummy variables and we allow for correlation within MSA. For model 10, the model includes MSA dummy variables and thus we allow correlation within state. The major change in table 38, as compared to table 37 is observed in model 10. Without allowing correlation within geographic unit, the Wald test shows that the difference in the coefficients of risk variables is significant (model 7). However, when we allow for correlation among geographic area, the difference is no longer significant. The coefficients of injury variables are still significantly different from each other at the 1% level in both model 9 and model 10.

When comparing model 9 and model 10, model 9 is preferred. Model 9 uses MSA-level violent and non-violent risk. The concern with model 10, which uses state-

[55] The complete set of models in table 38 with allowing the correlation among a same geographic unit are available in appendix G.

level violent and non-violent risk rates, is that the risk rates may not reflect driver's actual fatal risk at work particularly for violent risk cases. In model 10, state-level risk is uniformly assigned to the workers regardless of workers' residency. For example, taxi drivers working in MSAs and taxi drivers working in small towns in a same state are assigned a same level of violent fatal risk, which is not reasonable to assume. On the other hand, model 9 does not include drivers who live outside of MSA, which make the sample of drivers more homogeneous in terms of the fatal risk level they face. It is more reasonable to assume that the drivers who live in a same MSA face a same level of fatal risk (which still varies by occupation), than to assume that drivers who live in a same state face a same level of fatal risk.

Still, uniformly assigning the MSA-level risk rates for all drivers may not reflect their actual risk. We also assign different geographic level risk rates for different drivers. Truck drivers are likely drive all over the state or inter-state, and sales driver and taxi drivers are more likely to work in the MSA. Bus drivers could work all over the state if they are inter-state bus drivers. However, bus drivers may also work within a MSA (e.g. public transit bus drivers). Table 39 presents the results when we assign different geographic level risk rates to different drivers. In model 11 and 12, we assign the state-level risk rate for both violent and non-violent events for truck drivers and assign the MSA-level risk rates for both violent and non-violent events for bus, sales and taxi drivers. The models also allow for correlation within geographic area as has been done in table 38. Model 11 includes MSA dummy variables and model 12 includes state dummy variables. In both models, violent risk is significantly positive at the 1% level and non-violent risk is not significantly different than zero. The difference in coefficients of two risk variables is significant at the 5% level.

In model 13 and 14, we assign state-level violent and non-violent risk rates for truck and bus drivers, and MSA-level violent and non-violent risk rates for sales and taxi drivers. Assigning the state level risk rates for bus drivers reduce the magnitude of the coefficient of violent fatal risk, and increase the magnitude of coefficient of non-violent fatal risk. This change makes the difference between two coefficients of risk variables insignificant in both model 13 and model 14. Throughout the models in table 39, violent injury risk has significant positive coefficient and non-violent injury risk has a non-significant coefficient. The difference in coefficients of these injury variables is significant in all models.

It is hard to discuss which model, model 11/12 or 13/14, is preferred. If the number of local bus drivers outweighs the number of inter-state bus drivers, then model 11/12 would be preferred, and vice versa. There is no information in the CPS indicating whether a bus driver works at the local or inter-state level. However, it would not be unreasonable to assume that the number of local bus drivers outweigh the number of inter-state bus drivers if it is the case that public transit bus industry at a local level is bigger than the inter-state level bus industry. However, this model may contain the problem due to the multicolinearity between fatal risk rates and geographic dummy variables.

Table 40 examines the sensitivity of results when we drop the observations with extreme levels of risks. The risk rates used in these models are MSA-level for both violent and non-violent events and models include state dummy variables and allow correlation among workers by MSA. Model 15 excludes the workers who live in a state in which the *state-level* violent risk rate is zero. This exclusion drops 1,572 observations. The difference between two fatal risk coefficients is still significant at the 5% level. However, the difference between two injury risk coefficients becomes insignificant. This change is difficult to explain, and requires further characterization of dropped observations in terms of injury risks.

In model 16, we drop the workers who face the MSA-level violent fatal rate higher than 12.7 (in 10,000), which is approximately the 95th percentile for violent fatal risk in our sample. This exclusion drops 142 observations, and affects the coefficients of fatal risk variables. The coefficient of violent fatal risk is no longer significant and the difference of coefficients between two fatal risks is also not significant. The exclusion of workers who face a high level of violent fatal risk does not affect the coefficients of injury risks. The coefficient of violent injury risk is still significant at the 1% level and the difference between two injury coefficients is significant at the 5% level.

Model 17 drops both workers who live in a state with zero violent fatal risk as well as workers who face a MSA-level violent fatal risk higher than 12.7 (in 10,000). The results for this model are similar to model 14. The coefficient of violent fatal risk is significant (at the 5% level) and the difference in the coefficients between two fatal risk variables is also significant (at the10% level). Neither of injury risk coefficients are significant, nor the difference in coefficients of these variables. In summary, the estimation results are sensitive to the exclusion of extreme cases of violent fatal risk level from the sample. However, in all models presented here, workers require higher compensation to accept marginal increase of violent risk level than marginal increase of non-violent risk level, either from fatal injury events or non-fatal injury events.

Lastly, table 41 examines the robustness of results to including MSA variables on model 9. The MSA variables included are the MSA-level annual unemployment rate, the MSA-level sales volume per capita in wholesales, retail, transportation, entertainment and food industries, and the MSA-level per capita vehicle miles traveled. The MSA-level annual unemployment rates vary by year and the sales volumes is in 1997 level (see chapter 3). Table 41 shows the coefficients for the fatal risk, injury risk, and MSA variables. The coefficients of other demographic variables are not reported for succinctness.[56] Wholesales and retail sales volume per capita have positive significant coefficients at the 1% and 10% level, respectively. Transportation sales volume per capita has a negative coefficient that significant at the 10% level.

The vehicle miles traveled per capita has a negative coefficient which is significant at the 5% level. The coefficient of VMT per capita is small due to the large size of VMT per capita unit as compared to the unit for the log wage. The mean VMT per capita is 9,310 miles per year with standard deviation of 2,348 miles per year.[57] The increase of one standard deviation from the mean VMT per capita will reduce the wage by 1.6%. Inclusion of the MSA variables slightly increases the coefficients for fatal risks, but does not affect the comparison between types of fatal risks or types of injury risks. The violent fatal risk has significantly larger coefficients than non-violent fatal risk at the 5% level and the same is true for the injury risk cases.

---

[56] Results for other variables are not affected by the inclusion of MSA variables.
[57] This mean value is not from the sample and is computed assuming the each MSA receiving equal weight.

Conclusions.

This is one of few studies that use a revealed preference method to test the transferability of the VSL across different policy contexts. Improved study design upon the previous revealed preference study enables us to articulate the individual's WTP for different types of risk. This study overcomes the issue of hypothetical bias that may pertain to previous studies based on surveys, and provide insights about how individual risk perceptions might play an important role in their actual decision making.

We create unique geographic-occupation-specific fatal and injury risk rates for occupational drivers. The sample of drivers is collected from the monthly CPS. Our estimation results generally show that occupational drivers require more compensation to accept a marginal increase in violent fatal risk rate as compared to a marginal increase in a non-violent fatal risk rate. Results also show that occupational drivers require more compensation to accept a marginal increase in violent injury risk as compared to a marginal increase in a non-violent injury risks. The estimates of the wage/fatal risk relationship are somewhat sensitive to the different geographic level at which the fatal risks are created and for the geographic level of dummy variables included in the model. On the other hand, the estimates of the wage/injury risk relationships are quite robust to these changes.

The sensitivity of the wage/fatal risk relationships are likely to come from the less appropriate model specifications. For example, combining fatal risks and dummy variables which are created at a same geographic level in the same model is likely to be inappropriate due to the concern of multicolinearity among these variables. In addition, the state-level fatal risk, especially for violent fatal risk, is not a preferred risk rate since it ignores the heterogeneity of workers in terms of the fatal risk level at work within a same state.

Among all model specifications, we prefer models with MSA-level fatal risks that include state dummy variables over other models for two reasons. The first reason is that this model contains less multicolinearity issue. The second reason is that the MSA-level fatal risks have a better representation of worker's actual fatal risk level than state-level fatal risks. When we focus on the estimation results from models with MSA-level violent and non-violent risks that include state dummy variables, our wage/fatal risk estimation results are robust to different model specifications. Allowing correlation among workers within a same state, as well as adding MSA-specific demographic variables does not affect the results. Our wage/injury risk results are also robust to different models specifications just mentioned above.

Both wage/fatal risk and wage/injury risk results are somewhat sensitive to dropping workers with extreme violent fatal risk values from the sample. When we drop individuals who work in states with zero violent fatal risk (at the state-level), we lose the significance of the violent injury risk coefficients, while we keep the significance of the violent fatal risk coefficient. When we drop observations in high ends of MSA-level violent fatal risk (greater than the 95th percentile for our sample), we lose the significance of violent fatal risk coefficients while we keep the significance in violent injury risk. When we drop both observations in state with zero violent fatal risk and high ends of MSA-level violent fatal risk, we have significant violent fatal risk coefficient but non-significant violent injury risk. These changes of significance levels are puzzling. However, in any cases, either fatal risk or injury risk estimations indicate that drivers require higher compensation to accept marginal increase in violent risk than non-violent risk.

The VSL for total risk in our sample in our base model (model 2) is $1.8 million and our preferred model (model 6) is $1.4 million.[58] This is compatible with the VSL estimated from the panel models in chapter 4. Although the type of workers included in this chapter and chapter 4 are quite different, we control worker heterogeneity in both studies, thus lending validity to the convergence of the VSL estimates between these two samples. This study indicates that the VSL estimate from total risk is not applicable if we want to evaluate risk-specific policies such as traffic risk or violent risk. When we separately evaluate individual's willingness to pay to reduce each type of risk, we find quite different point estimates for WTP. We find much higher VSL based on violent risk coefficient, and we do not find any wage compensation toward non-violent risk.

The VSL estimated for violent risk in this study (around $3-5million) is in a range of the VSL estimated from previous cross section HW models, that is $4-10 million (Viscusi 1993). At first thought, this comparison may appear to be counter to our hypothesis that wage/risk premia are differentiated by type of risk. However, as discussed in chapter 4, existing estimates of the VSL in the literature typically do not control for unobserved heterogeneity. In chapter 4, when we use undifferentiated risks and control for unobserved heterogeneity, our estimates of the VSL are approximately $2 million, lower than the $4-10 million range found in the comparable existing literature using undifferentiated risks.[59]

---

[58] The coefficient of total risk in model 6 is 0.0044.
[59] Total risk is the sum of violent and non-violent risks. The wage/risk premium for total risk is likely smaller than that for violent risk because we could not find a significant compensation for non-violent risk, suggesting this type of risk may not be considered an issue for occupational drivers (perhaps they feel it is

We argue in chapter 4 that $4-10 million for undifferentiated risks is likely an upwardly biased estimate. Our results in this chapter for undifferentiated risks also suggest $4-10 million is upwardly biased. As such, our results suggest that a VSL of around $4 million may be appropriate to use in policy analysis designed to reduce the risk of fatal violent injuries among occupational drivers, and would not be appropriate to use for other policies involving different types of risks. This point is further underscored by the fact that we could not find a significant wage premium for increased traffic accident risk among occupational drivers. Traffic accident risk may be perceived as a highly controllable risk by this group of individuals, and as such, may not require wage compensation for working in areas with higher levels of fatality risks. How the VSL level changes when we change the target risk type or target worker population should be pursued in the future study.

The value of statistical injury for total injury risk in our sample in our base model (model 3) is $70,000 and our preferred model (model 6) is about $52,000.[60] This is in line of previous estimated value of statistical injury. However, once we separate violent and non-violent injury risk, we find much higher value of statistical injury for violent injury risk, which is around $1 million while we do not find significant wage compensation towards non-fatal non-violent injury risk. This result confirms the conclusions from fatal risk analysis.

Although the focus is on the individual's behavior within the labor market, this study is potentially important to evaluate the transferability of VSL across different policy contexts. Our results indicate that individuals exhibit different WTP to avoid

---

in their control). As such, when adding violent and non-violent risks together to get a total risk, we are adding measurement error to the "risk" of the job, biasing this coefficient towards zero.

[60] The coefficient of total injury risk in model 6 is 1.4544.

different types of risks even in the same policy context. This makes it hard to justify the application of VSL transfer between different policy contexts. Although the occupational drivers are a special sample, we show that transferring VSL from one context to another is not an appropriate policy evaluation approach. Our results from the revealed preference method confirm the previous results from stated preference method, and suggest risks with more dread/fear or less controllability require more compensation than risks with less dread/fear or more controllability. These results suggest that it is necessary to estimate VSLs in the same policy context, or that any benefits transfer exercise require close attention to the types of risk being evaluated.

Table 35

Regression Result: (Dependent Variable: ln(weekly wage), without MSA<100)

	Model 1 Risk data 1992-2002		Model 2 Risk data 1998-2002	
	Coefficient[a]	Standard error	Coefficient[a]	Standard error
Violent risk (MSA level)	0.0050***	0.0016	0.0140***	0.0042
Non-violent risk (state level)	-0.0087	0.0063	-0.0138**	0.0058
age	0.0373***	0.0022	0.0369***	0.0024
agesq	-0.0003***	0.00002	-0.0003***	0.00002
ugdeg	0.00002***	0.0195	0.1011***	0.0218
college	0.1012***	0.0109	0.1032***	0.0122
hsgrad	0.1000***	0.0096	0.0962***	0.0108
hispanic	-0.0745***	0.0137	-0.0701***	0.0153
blacknh	-0.0348***	0.0114	-0.0222*	0.0126
othrace	-0.0978***	0.0256	-0.0704**	0.0290
uscit	0.1063***	0.0157	0.1273***	0.0173
female	-0.1800***	0.0153	-0.1754***	0.0168
salary	0.0833***	0.0080	0.0900***	0.0089
workot	0.2716***	0.0073	0.2718***	0.0082
union	0.2341***	0.0082	0.2188***	0.0092
married	0.0741***	0.0079	0.0696***	0.0088
central	-0.0276***	0.0089	-0.0362***	0.0101
truck	0.0675***	0.0207	0.0860***	0.0220
bus	-0.0245	0.0205	-0.0128	0.0234
taxi	-0.1344***	0.0290	-0.1480***	0.0340
Neweng	-0.0389	0.2815	-0.0975	0.1930
Midalt	-0.3566	0.2899	-0.1698	0.1309
Encent	0.0639	0.2766	-0.0384	0.2918
Wncnet	0.1350	0.2852	0.0497	0.3006
Satl	0.1601	0.3030	0.0598	0.2575
Escent	0.1833	0.2833	0.1032	0.2838
Wscent	-0.5688*	0.3260	-0.1920	0.1589
Mount	-0.3271	0.2928	0.1386	0.1493
constant	5.5808***	0.2803	5.2821***	0.1076
Test: violent= non-violent (p-value)	0.05		0.0004	
N	12,637		10,257	
Average weekly wage ($)	695.79		701.65	
VSL (violent risk)	$1.8 million		$5.1 million	
R2	0.32		0.32	

Note. MSA and year dummies are omitted to report.
[a] regression shows significant heteroschedasity at 5% level and thus use robust estimators.
***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

Table 36

Results With MSA Level Violent Risk and State Level Non-violent Risk Without MSA<100 (Standard error in parentheses)

	Model3: w/ MSA dummy	Model4: w/ state dummy
violent fatal risk (MSA level)	0.0121***	0.0121***
non-violent fatal risk (state level)	-0.0043	-0.0028
violent injury risk	29.2671***	30.8813***
non-violent injury risk	0.5831	0.1285
average weekly wage ($2004)	697.64	697.64
R2	0.32	0.30
N	8,685	8,685
fatal: violent=non-violent (p-value)	0.08	0.102
non-fatal: violent=non-violent (p-value)	0.008	0.0044
VSL (violent fatal risk)	$4.3 million	$4.3 million

*p<0.1, **p<0.05, ***p<0.01

Table 37

Results With MSA Level or State level Fatal Risks Without MSA<100

(Standard error in parentheses)

	MSA-level fatal risks		State-level fatal risks	
	Model 5: w/ MSA dummy	Model 6: w/ state dummy	Model 7: w/ MSA dummy	Model 8: w/ State dummy
violent fatal risk	0.0110***	0.0115***	0.0150**	0.0136**
non-violent fatal risk	0.0010	0.0003	-0.0001	0.0084
violent injury risk	29.5636***	31.0635***	38.1399***	29.8873***
non-violent injury risk	0.7203	0.2158	0.2217	0.8323
average weekly wage ($2004)	697.64	697.64	689.69	689.69
R2	0.32	0.30	0.30	0.29
N	8,685	8,685	12,509	12,509
fatal: violent=non-violent (p-value)	0.134	0.0256	0.0328	0.528
non-fatal: violent=non-violent (p-value)	0.0085	0.0043	0.0001	0.0042
VSL (violent fatal risk)	$3.9 million	$4.1 million	$5.3 million	$4.8 million

*p<0.1, **p<0.05, ***p<0.01

Table 38

Results with Allowing Correlation Within a Geographic Unit (standard error in parentheses)

	MSA-level fatal risks	State-level fatal risks
	Model 9: w/ State dummy	Model 10: w/ MSA dummy
violent fatal risk	0.0115***	0.0150*
non-violent fatal risk	0.0003	-0.0001
violent injury risk	31.0635**	38.1399***
non-violent injury risk	0.2158	0.2217
average weekly wage ($2004)	695.28	687.31
R2	0.30	0.30
N	8,685	12,509
fatal: violent=non-violent (p-value)	0.0131	0.1999
non-fatal: violent=non-violent (p-value)	0.0215	0.0009
VSL (violent fatal risk)	$4.1 million	$5.3 million

*p<00.1, **p<00.05, **p<00.01

Table 39

Different Geographic Risk Level by Type of Driver With Allowing Correlation Within a Geographic Unit (standard error in parentheses)

	truck: state-level, others: MSA-level fatal risks		Truck/bus: state-level, others: MSA-level fatal risks	
	Model 11: w/ MSA dummy	Model 12: w/ State dummy	Model 13: w/ MSA dummy	Model 14 w/ State dummy
violent fatal risk	0.0127***	0.0119***	0.0104***	0.0082**
non-violent fatal risk	-0.0047	-0.0018	0.0005	0.0087*
violent injury risk	34.6514***	26.8280*	42.3898***	33.8273**
non-violent injury risk	-0.1886	-0.0503	-0.0701	0.4533
average weekly wage ($2004)	691.40	691.40	688.82	688.82
R2	0.30	0.29	0.30	0.29
N	12,131	12,131	12,291	12,291
fatal: violent=non-violent (p-value)	0.0119	0.0358	0.1796	0.9397
non-fatal: violent= non-violent (p-value)	0.0003	0.0537	0.0001	0.0179
VSL (violent fatal risk)	$4.5 million	$4.2 million	$3.7 million	$2.9 million

*p<0.1, **p<0.05, ***p<0.01

Table 40

Results With Limited Risk Range (With State Dummy and Allowing Correlation Within MSA) (standard error in parentheses)

	Model 15 Drop vrisk_s=0	Model 16 Drop vrisk_m>12.7	Model 17 Drop vrisk_s<0 & vrisk_m>12.7
violent fatal risk (MSA level)	0.0139***	0.0105	0.0163**
non-violent fatal risk (MSA level)	0.0021	0.00003	0.0018
violent injury risk	24.4766	30.6181**	23.2210
non-violent injury risk	0.6544	0.2835	0.8628
average weekly wage ($2004)	696.74	698.90	698.26
R2	0.32	0.33	0.31
N	7,113	8,514	6,942
fatal: violent=non-violent (p-value)	0.0135	0.16	0.08
non-fatal: violent=non-violent (p-value)	0.1480	0.0244	0.18
VSL (violent fatal risk)	$5.0 million	$3.8 million	$5.9 million

*p<0.1, **p<0.05, ***p<0.01

Table 41

Results With MSA Variables With Allowing Correlation Within MSA (standard error in parentheses)

	Model 18: w/State dummy
violent fatal risk (MSA-level)	0.0133***
non-violent fatal risk (MSA-level)	0.0021
violent injury risk	30.9522**
non-violent injury risk	-0.2817
unemp	-0.0043
whole_sales	0.0035***
retail_sales	0.0063*
trans_sales	-0.0195*
ent_sales	-0.0225
food_sales	0.0082
msavmtp	-7.06e-06**
average weekly wage ($2004)	692.43
R2	0.31
N	7,082
fatal: violent=non-violent (p-value)	0.0388
non-fatal: violent=non-violent (p-value)	0.0329
VSL (violent fatal risk)	$4.7 million

*p<0.1, **p<0.05, ***p<0.01

Chapter VI

Conclusions

This dissertation addressed two important issues in the VSL literature. The first issue is the potential endogeneity bias in cross-section hedonic wage models. The second issue is the transferability of the VSL between different policy contexts.

Chapter 4 addressed the issue of endogneity bias in cross-section HW models. We first estimated the cross-section model and panel models to identify the bias due to the time-invariant worker heterogeneity. We also combined panel models and instrumental variable approach to control potential remaining endogeneity bias due to the measurement error associated with risk variable, time-variant worker heterogeneity and simultaneity between wage and risk.

We use the national panel data of Survey of Income and Program Participation as our labor market data and occupation-industry risk matrices from Scotton (2000) as occupation-industry fatal risk data. We find a VSL of $4.6 million (in 2005 dollars) with a standard error of $0.6 million from the cross-section hedonic wage model. After controlling for workers unobserved time-invariant heterogeneity with the fixed-effect and first-difference models, we find the VSL of $2.5 million with the standard error of $0.4 million and the VSL of $1.7 million with the standard error of $0.6 million, respectively. With the 95 percent confidence interval, there is no overlap of the VSL estimated from the cross-section model and panel models, while there is overlap the VSL estimated from the fixed-effect model and first-difference model. We find no evidence of exacerbated attenuation bias from measurement error or remaining endogeneity bias in our panel models. We conclude that the cross-section OLS hedonic wage model is significantly biased upward due to the unobserved time-invariant worker heterogeneity, but not from the time-variant worker heterogeneity or simultaneity between wage and risk.

Our hedonic wage models are sensitive to the inclusion of industry dummy variables and firm-side variables. When we do not include industry dummy variables, our VSL estimate from the cross-section model increases to $10.3 million. When we further drop firm-side variables, the VSL estimate from the cross-section model decreases to $7.2 million. As documented in previous studies (McConnell, 2006; Mrozek & Taylor, 2002), we find that controlling industry differences and firm-side characteristics are important to obtain unbiased VSL estimates.

The inter-industry wage differentials are well documented phenomenon in labor economics, and theoretically, industry dummy variables should be included in the hedonic wage model (McConnell, 2006). The reason for not including industry dummy variables in hedonic wage models often is stated as a concern for potential multicolinerity between the risk variable and the industry dummy variables (Viscusi & Aldy, 2003). This concern is valid when fatal risk data only varies by industry. Including the industry dummy variables often removes large amount of variation in risk variables and generate insignificant or sometimes negative risk coefficient (Dorman & Hagstrom, 1998).

Our risk data in which risk rates vary by occupation and industry should mitigate the problem of multicolinerity between risk and industry variables. Our regression results show that omitting industry variables significantly bias the risk estimators. Therefore when the industry dummies are omitted, the model may be significantly biased and that must be corrected with instrumental variables models.[61] Our regression models employ the one-digit level industry variables which are more aggregated than the industry variation in the risk variable. Our second stage panel estimation results indicate that there is not an endogeneity problem in our panel models due to not controlling for more detailed level inter-industry wage difference.

Mrozek and Taylor (2002) suggests that the VSL obtained from previous cross-section hedonic wage models assuming all studies include industry dummy variables is about $2 million dollars in 1998 dollars or $2.4 million in 2005 dollars. This is a quite similar value to the VSL we obtained from our panel models. This may be because previous cross-section hedonic wage studies suffer from two types of bias; omitted variables bias and measurement error bias associated with risk variable, which work in opposite directions in this case. Our study indicates that the previous studies suffer from upward bias due to the omitted variables. However, at same time, previous cross-section hedonic wage studies are likely suffered from the attenuation bias from measurement error associated with risk variables.

Since measurement error bias has been one of the major concerns in hedonic wage literature (Viscusi & Aldy, 2003), we further discuss this issue. Most studies considered in Mrozek and Taylor (2002) use industry or occupation average risk data. When studies assign industry average risk levels to workers, they ignore the variation of risk within the industry. In the same way, when studies assign the occupation average risk levels to workers, they ignore the variation of risk within the occupation. For example, secretaries and construction workers in the construction industry must face different levels of risk. However, in the hedonic wage model, researchers have had to

[61] However, exclusion of industry dummy variables may lead to poor fit of the first stage regression.

assign the same risk level for these two types of workers due to data limitations. This creates a significant measurement error problem in hedonic wage models and biases risk estimators downward. In our hedonic model, we assign risk rates that vary by occupation and industry. Thus in our example, secretaries and construction workers in the construction industry are assigned different risk levels. Thus our study might mitigate measurement error bias caused from the disparity between *actual* risk and *estimated* risk level.

Although our study mitigates the measurement error bias caused from the disparity between actual risk and estimated risk level, there is still a potential measurement error problem arising from the disparity between *actual* risk and *perceived* risk. It is reasonable to assume that the workers do not know their exact actual risk level (such as 3.5 in 100,000 chance of death), but have some perceived risk level which may be different from the actual risk level. A major concern is that this measurement error may exacerbate attenuation bias in panel models, which were estimated in chapter 4. To attempt to address this issue, two stage panel estimations were employed.

It was somewhat surprising to find that there was not measurement error bias in our panel models. There are two potential reasons why measurement error was not an issue in our hedonic wage model. The first reason is that the measurement error we were concerned with may not satisfy the classical errors in variables (CEV) assumption. The CEV assumption indicates that measurement error biases the estimator only if there is a correlation between measurement error and the objective (actual) risk level. If this is not the case, then the measurement error only increases the variance, and would not cause any bias in the estimator (Wooldridge, 1999).

The other reason is that the labor market may be in a long-run equilibrium and firms have correct information about the actual risk level. As discussed in chapter 2, if the market is in a long-run equilibrium, the hedonic wage schedule is solely determined by the distribution of firm' isoprofit functions. If firms perceive the risk level in the same way as researchers do (i.e., perceive objective risks accurately), then a hedonic wage function is unbiased even if workers perceive risks differently from the actual risks.

As compared with Kniesner et al.(2005), even after adjusting the difference of model specifications, the VSL estimates from our panel models are lower. When we omit industry and firm side variables, we obtain the VSL of $3.0 (in 2005 dollars), while Kniesner et al. (2005) obtain the VSL of $6.7 million,[62] both from the first difference model. This is a quite large difference that we should not ignore. The reason of the disparity between our VSL estimates and Kniesner et al.'s (2005) estimates may be the difference of our labor market data. We use the SIPP as our labor market data while Kniesner et al. use the PSID as their labor market data. According to the SIPP User Guide, neither the SIPP nor the PSID are designed to gather a representative sample within each state.[63] This indicates that the distribution of workers characteristics may be significantly different between the SIPP and the PSID. In addition, the PSID data set tends to generate high-end VSL estimates. According to Viscusi and Aldy (2003), the VSL estimated using the PSID is between $8-20 million while the VSL estimated using the CPS is $0.7-12 million (in 2000 dollars).

In the future analysis, we should include the CPS as the labor market data to examine the robustness of our results. The CPS is designed to be a representative

[62] Kniesner et al. (2005) do not explicitly mention to which year they adjust their dollar values. Here we assume they use 2005 dollar value.
[63] Both SIPP and PSID are designed to over-sample low-income population.

sample of households within each state.[64] Therefore the panel estimates from the CPS would have more relevance to be used in the policy analyses. Although the CPS does not provide enough firm-side variables, we can estimate the possible bias due to the lack of firm-side variables from this study. In addition, the CPS only has two current wage observations for each individual, which may limit the sample size. Therefore, the results of this study and the results from future study using the CPS should be combined to provide an overall assessment of VSL estimates arising from current risk and labor market data.

Chapter 5 addressed the issue of the transferability of the VSL between different policy contexts using cross-section hedonic wage models. We examined whether or not workers and firms differentiate heterogeneous risks to determine the risk-wage compensation levels. We focus on two very different fatal risks in terms of the degree of workers' control over the risk and the degree of dread associated with risk. We use risks related to violent assaults and risks related to non-violent events. We use occupational drivers to mitigate potential unobserved heterogeneity of job characteristics and measurement error associated with risk variables. The labor market data comes from the basic CPS, and the occupation-geographic specific risk rates for each cause of death are created from the non-public Census of Fatal Occupational Injuries.

When we use the MSA level risk rates for both violent and non-violent fatal risks, we find quite robust evidence of worker's different WTP to reduce a marginal risk for different types of risks. We find that occupational drivers require larger compensation to accept a marginal increase of violent risk as compared to non-violent risk. This is true for both fatal and non-fatal risk cases. When we use different geographic combinations

[64] According to the SIPP User Guide.

of violent and non-violent risk rates, the results are less robust for the changes in the model specifications, but this is largely due to inappropriate model specifications. The injury risk data is only available at the state level, and the estimations of injury-risk wage compensation are quite robust to changes in model specifications.

This study verifies the findings from previous contingent valuation studies regarding the WTP to reduce heterogeneous risks. The contribution of this study is that we show that individuals exhibit different WTP to reduce different types of risks by a revealed preference method. The contingent valuation method is a very flexible and useful tool in many settings, however this method may suffer from hypothetical bias due to the nature of the surveys. The revealed preference (hedonic wage) method, on the other hand, faces certain limitations due to the data availability. Slovic et al (1980) indicates that violent risk is less controllable and more dreadful than traffic accident risk, which is the major component fatal risk for occupational drivers. Our results strongly support that occupational drivers and firms differentiate heterogeneous risks depending of the qualitative risk characteristics, and generate different hedonic wage functions for each type of risk.

Although both risks we examined in this study are occupational risks, we can derive the implication of transferability of the VSL between different policy contexts. If individuals differentiate heterogeneous types of risks in a similar circumstance (such as risk at work), it is hard to assume that individuals do not differential heterogeneous risks that arise in very different circumstances, such as risk at work and risk from environmental damages. Thus this study suggests that current direct use of VSL obtained from hedonic wage studies in benefit estimation of various governmental programs should be reconsidered.

Our study does not aim to quantify the difference of worker's WTP for the occupational risk and other types of risks. Instead, we aim to examine whether or not individuals differentiate risks with very different risk characteristics in their risk-wage compensation decision using the revealed preference method. In the future analysis, we would incorporate more variety of risks such as exposure to harmful substances or environment and examine the sensitivity of workers risk preference towards other types of risks.

To provide more direct implications to the environmental policy analysis, we would want to quantify the difference of worker's WTP to reduce environmental-related risk and other risks. Although there is no identical risk as environmental risk available in occupational setting, the comparison between worker's WTP to reduce exposure to harmful substances or environment and other risks at work may provide more direct implications to assess the validity of current environmental policy analysis.

In future analyses, it also would be important to examine robustness of our results with different groups of workers. The fact that none of our estimation results show significant coefficients for non-violent fatal or injury risk may suggest that there is not enough variation in these variables to estimate wage/risk (wage/injury risk) coefficients among drivers. Or, it could be the case that risks with great perceived personal control such as auto accident risks, the primary component of occupational driver's non-violent risk, are not compensated in the workplace. Adding different types of workers may enable us to estimate the non-violent risk premium better, since it would add workers who face different traffic-related risk level than occupational drivers, which would increase the variation in non-violent risk/injury variable. In addition, non-occupational driving workers may have different perceived personal control over the traffic-related risk, and they may require a different compensation over the traffic-related risk as compared to occupational drivers.

To include different types of workers in the model, it is important to control for time-invariant worker heterogeneity as found in chapter 4. In chapter 5, we used workers with similar job requirements to control unobserved job as well as worker characteristics. However, if we include more variety of workers, we would have same problems related to the unobserved worker heterogeneity as in chapter 4. We can solve this omitted variable problem by employing the panel data analysis as in chapter 4.

Appendix A CFOI Research File Elements[a]

Element	Description
REC	Record ID
REF	Reference year
YEA	Year of injury
MON	Month of injury
DAY	Day of week
TII	Time of incident
NAT	Nature of injury
PAR	Part of body
EVE	Event or exposure
SOU	Source of injury
SEC	Secondary source of injury
ACT	Worker activity
LOC	Location
REG	Region
OCC	Occupation
USO	Usual lifetime occupation
IND	Industry
USI	Usual lifetime industry
OWN	Ownership
EST	Establishment size class
EMP	Employee status
TIE	Time with employer
GEN	Gender
AGE	Age group
RAC	Race
HIS	Hispanic origin
FOR	Foreign born
SUR	Days survived
NARR	Narrative

[a] Source: BLS, CFOI manual.

Appendix B Industry Group Classification[a]

Major industry group	Dummy name	SIC Division	Two-digit SIC codes
Agriculture, Forestry, Fishing and Mining	AGIND	Division A and B (10 and 20)	01-14
Construction	CONSTIND	Division C (30)	15-17
Manufacturing	MANUFIND	Division D (41: Durable and 42: Non-Durable)	20-39
Transportation, Communications, Electric, Gas, and Sanitary Services	TCUIND	Division E (51: Transportation and 52: Communication & Utilities)	40-49; except 43
Wholesale trade and Retail trade	TRDIND	Divisions F and G (61: Wholesale and 62: Retail)	50-59
Finance, Insurance, Real Estate, and Services	SERVIND	Divisions H and I (70 and 80)	60-88
Public Administration	PUBIND	Division J (90)	43, and 91-99

[a] Reproduced from Scotton (2000), pp.193.

Appendix C Occupation Group Classification[a]

Major occupation group	Dummy name	Sub-group	CPS Codes
Managerial and professional specialty occupations	PROFOCC	Managerial	003-037
		Professional	043-199
Technical, sales, and administrative support occupations	TECHOCC	Technicians and related support occupations	203-235
		Sales occupations	243-285
		Administrative support occupations, including clerical	303-389
Service occupations	SERVOCC	Private household	403-407
		Protective service	413-427
		All other service occupations	433-469
Precision production, craft, and repair occupations	CRAFTOCC		503-699
Operators, fabricators, and laborers	LABOROCC	Machine operators, assemblers, an inspectors	703-799
		Transportation and material moving equipment occupations	803-859
		Handlers, equipment cleaners, helpers and laborers	864-889
Farming, forestry, and fishing occupations	FARMOCC		473-499
Armed forces and unidentified	(not used)		

[a] Reproduced from Scotton (2000), pp.199.

Appendix D Occupation Group Mapping[a]

Occ code	22 Occupation Groups	Census Occupation Classification codes
70120	Executive & Administrative Positions	003-022
70300	Management Related Occupations	023-037
70400	Engineers	044-059
71290	Professional Occupations (except Engineers)	043, 063-066, 069, 73-79, 083-106, 113-199
71590	Technicians (includes air craft pilots)	203-235
71900	Marketing and Sales Occupations	243-285
72300	Secretaries & Typists	313-315
72400	Financial Records Keepers	337-344
72600	Administrative Support Occupations (except Finance & Secretaries)	303-309, 316-336, 345-389
73100	Cleaning & Building Service and Maintenance	448-455
73200	Service Workers (except Cleaning & Building Service and Maintenance)	404-447, 456-469, 425-432
73350	Mechanics (all types)	505-552
73400	Blue-Collar Worker Supervisors	503, 553-558, 613, 628, 803, 843, 864
73490	Construction Tradesmen	563-599
73510	Extractive Occupations	614-617
73540	Precision Workers	634-699, 796-799
73630	Machine Operators	703-779, 796-799
73700	Fabricators & Hand workers	783-795
73820	Truck Drivers	804
73900	Motor Vehicle & Material Moving Equip Operators	806, 808, 813-814, 823-859
74000	General Laborers	865-889
74390	Farming, Forestry & Fishing Occupations	473-499

[a]Source: Scotton (2000) pp.200 with some corrections.

Appendix E Industry Group Mapping[a]

23 Industry Groups	23 Inds Code	Industry (2-digit SIC code)	SIC	SIPP code
Agriculture, Forestry, and Fishing	9010	Agricultural Production Crops	01	010-032
	9010	Agricultural Production Livestock and Animal Specialties	02	
	9010	Agricultural Services	07	
	9010	Forestry	08	
	9010	Fishing, Hunting, and Trapping	09	
Mining, Extraction and Quarrying	9020	Metal Mining	10	040-050
	9020	Coal Mining	12	
	9020	Oil and Gas Extraction	13	
	9020	Mining and Quarrying of Nonmetallic Minerals, Except Fuels	14	
Construction	9030	Building Construction General Contractors and Operative Builders	15	060
	9030	Heavy Construction Other Than Building Construction Contractors	16	
	9030	Construction Special Trade Contractors	17	
Food and Tobacco Products	9420	Food and Kindred Products	20	100-130
	9420	Tobacco Products	21	
Textile Mill and Apparel Products	9423	Textile Mill Products	22	132-152
	9423	Apparel and Other Finished Products from Fabrics & Similar Materials	23	
Lumber/Wood/Stone/Glass Products	9432	Lumber and Wood Products, Except Furniture	24	230-262
	9432	Furniture and Fixtures	25	
	9432	Stone, Clay, Glass and Concrete Products	32	
Paper and Printing Products	9427	Paper and Allied Products	26	160-172
	9427	Printing, Publishing, and Allied Industries	27	

Chemicals/Petro/Plastics/Leather Goods	9431	Chemicals and Allied Products	28	180-222
	9431	Petroleum Refining and Related Industries	29	
	9431	Rubber and Miscellaneous Plastics Products	30	
	9431	Leather and Leather Products	31	
Metals, Machinery, and Misc. Manufacturing Industries	9435	Primary Metal Industries	33	270-350
	9435	Fabricated Metal Products, Except Machinery & Transportation Equipment	34	
	9435	Industrial and Commercial Machinery and Computer Equipment	35	
	9435	Electronic & Other Electrical Equipment, Components, Except Computer Equipment	36	
	9435	Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods; Watches and Clocks	38	
	9435	Miscellaneous Manufacturing Industries	39	
Motor Vehicle and Equipment Manufacturing	9437	Transportation Equipment	39	351-370
Railroad and Water Transportation	9500	Railroad Transportation	40	400, 420
	9500	Water Transportation	44	
Personal Transportation Services (ground)	9541	Local/Suburban Transit & Interurban Highway Passenger	41	401, 402, 432
	9541	Transportation Services	47	
Trucking, Warehousing and Air Transportation	9545	Motor Freight Transportation and Warehousing	42	410-411, 421
	9545	Transportation by Air	45	
Communications, utilities and Sanitary Services	9549	Communications	48	440-442, 450-472
	9549	Electric, Gas, and Sanitary Services	49	
	9549	Pipelines, Except Natural Gas	46	

Wholesale Trade	9651	Wholesale Trade-durable Goods	50	500-574
	9651	Wholesale Trade-non-durable Goods	51	
Retail Trade	9652	Building Materials, Hardware, Garden Supply and Mobile Home Dealers	52	580-694
	9652	General Merchandise Stores	53	
	9652	Food Stores	54	
	9652	Automotive Dealers and Gasoline Service Stations	55	
	9652	Apparel and Accessory Stores	56	
	9652	Eating and Drinking Places	58	
	9652	Miscellaneous Retail (Liquor and Drug Stores)	59	
Finance, Insurance and Real Estate	9760	Depository Institutions	60	700-714
	9760	Non-depository Credit Institutions	61	
	9760	Insurance Carriers	63	
	9760	Insurance Agents, Brokers and Service	64	
	9760	Real Estate	65	
	9760	Holding and Other Investment Offices	67	
Personal Services	9872	Personal Services	72	761, 771-795
	9872	Private Households	88	
Business, Auto and Repair Services	9876	Business Services	73	721-760, 882-893
	9876	Automotive Repair, Services and Parking	75	
	9876	Miscellaneous Repair Services	76	
	9876	Engineering, Accounting, Research, Management and Related Services	87	
Entertainment Services	9879	Motion Pictures	78	800-810
	9879	Amusement and Recreation Services	79	
Health Services	9880	Health Services	80	812-840

Social, Legal, Educational and Other Services	9885	Hotels, Rooming Houses, Camps, and Other Lodging Places	70	
	9885	Legal Services	81	
	9885	Educational Services	82	762-770, 841-881
	9885	Social Services	83	
	9885	Museums, Art Galleries, and Botanical and Zoological Gardens	84	
	9885	Membership Organizations	86	
Public Administration & USPS	9990	United States Postal Service	43	412, 900-932
		All Other Public Administration	91-99	

[a]Source: Scotton (2000) p.194-198 with some modification.

Appendix F Definition of Occupational Drivers.

Occupation title	COC code	SOC 1980 code	Standard Occupation Classification 1980 definition	Occupation code [for 1998 (for 1999-2003)]and definition category for Occupational Employment Statistics
Truck driver	804	8212	Truck drivers, tractor-trailer includes semi-tractor and trailer truckers	97120 (53-3032): Truck drivers, heavy or tractor-trailer. Drive a tractor-trailer combination or a truck with a capacity of at least 3 tons, to transport and deliver goods, livestock or materials in liquid, loose or packaged form. May be required to unload truck.
		8213	Truck drivers, heavy, single body trucks of at least three tons weight, including dump, flat bed, redi-mix, tank trucks, and trucks mounted with special service equipment as tow trucks, etc.	
		8214	Truck drivers, light (including delivery and route drivers) operating automotive trucks less than 3 tons weight, including pick-up, delivery, and van trucks.	97105 (53-3033) Truck drivers, light include delivery and route workers. Driver a truck, van, or automobile with a capacity under 3 tons. May driver light truck to deliver or pick up merchandise. May load and unload truck

Appendix F (continued).

Sales driver	806	8218	Driver-sales workers, includes occupations concerned with driving trucks or other vehicles over established routes to deliver and sell goods such as bakery and dairy products; collect and deliver items such as laundry and dry-cleaned garments; or collect coins, refill vending machines and service vending machines	97117: Driver/ Sales Workers. Driver truck or other vehicle over established routes to: deliver and sell goods, such as food products; pick up and deliver items, such as laundry; or refill and collect coins from vending machines. Include newspaper delivery drivers.
Bus driver	808	8215	Bus drivers, includes occupations involving transporting passengers by bus including school, inter and intra city, and private.	97108 (53-3021): Bus drivers. Drive bus, transporting passengers over specified routes to local or distant points according to a time schedule. Assist passengers with baggage. Collect tickets or cash fares. 97111: Bus drivers, school: transport students between pick-up points and school. Maintain order during trip and adhere to safety rules when loading and unloading pupils.
Taxi driver	809	8216	Taxicab drivers and chauffeurs, includes occupations involving operating automobiles, limousines, and hearses to transport passengers and merchandise and driving new automobiles between production and customer delivery.	97114 (53-3041): Taxi drivers and chauffeurs. Driver automobiles, limousines, custom-built sedans, or hearses to transport passengers or cargo. May drive automobiles for delivery. Exclude ambulance drivers and bus drivers.

[a] Source: BLS, Occupational Employment Statistics, Dictionary of Occupations Customer Copy 1997-1998 retrieved April 25, 2007 from ftp://ftp.bls.gov/pub/special.requests/oes/oesdic_98.pdf.

Appendix G Complete set of results with allowing correlation within a geographic unit (Standard error in parentheses)

	MSA-level fatal risks		State-level fatal risks	
	Model9: w/ MSA dummy	Model10: w/ State dummy	Model11: w/ MSA dummy	Model 12 w/ State dummy
violent fatal risk	0.0110***	0.0115***	0.0150*	0.0136**
non-violent fatal risk	0.0010	0.0003	-0.0001	0.0084*
violent injury risk	29.5636***	31.0635**	38.1399***	29.8873**
non-violent injury risk	0.7203	0.2158	0.2217	0.8323
average wage ($2004)	695.28	695.28	687.31	687.31
R2	0.32	0.30	0.30	0.29
N	8,685	8,685	12,509	12,509
fatal: violent=non-violent	0.1378	0.0131	0.1999	0.4854
non-fatal: violent=non-violent	0.0041	0.0215	0.0009	0.0275
VSL (violent fatal risk)	$3.9 million	$4.1 million	$5.3 million	$4.8 million

*p<0.1, **p<0.05, ***p<0.01

References

Arabsheibani, G. R., & Marin, A. (2001). Self-selectivity bias with a continuous variable: Potential pitfall in a common procedure. *Applied Economics, 33*, 1903-1910.

Baum, C. F. (2006). *An Introduction to Modern Econometrics Using Stata*: Stata Press.

Benjamin, D. K. (2001). Individuals' estimates of the risks of death: Part II-new evidence. *Journal of Risk and Uncertainty, 22*(1), 35-57.

Black, D. A., Galdo, J., & Lin, L. (2003). *How robust are Hedonic wage estimates of the price of risk?* : Environmental Protection Agency, Office of Policy, Economics and Innovation, National Center for Environmental Economics.

Bluestone, P. (Forthcoming). *Government fragmentation and the attainment of regional environmental quality.* Georgia State University, Atlanta.

Brown, C. (1980). Equalizing differences in the labor market. *The Quarterly Journal of Economics, 94*(1), 113-134.

Card, D. (1999). The causal effect of education on earnings. In O.Ashenfelter & D.Card (Eds.), *Handbook of Labor Economics* (Vol. 3, pp. 1801-1859): Elsevier Science B.V.

Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., et al. (2002). Public perceptions of risk and preference-based values of safety. *Journal of Risk and Uncertainty, 25*(3), 211-232.

Combs, B., & Slovic, P. (1979). Causes of death: Biased newspaper coverage and biased judgements. *Journalism Quarterly, 56*, 837-843, 849.

Cookson, R. (2000). Incorporating psychosocial considerations into health valuation: An experimental study. *Journal of Health Economics, 19*(3), 369-401.

Corso, P. S., Hammitt, J. K., & Graham, J. D. (2001). Valuing mortality-risk reduction: Using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty, 23*(2), 165-184.

Cropper, M. L., & Subramanian, U. (1999). Public choices between lifesaving programs: How important are lives saved? In M. Cropper (Ed.), *Valuing Environmental Benefit*: Edward Elgar.

Cummings, R., & Taylor, L. O. (1999). Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. *American Economic Review, 83*(3), 649-665.

Dilingham, A. E. (1985). The influence of risk variable definition on value of life estimates. *Economic Inquiry 24*(April), 277-294.

Dorman, P., & Hagstrom, P. (1998). Wage compensation for dangerous work revisited. *Industrial and Labor Relations Review, 52*(1), 116-135.

Fisher, A., Chestnut, L. G., & Violette, D. M. (1989). The value of reducing risks of death: A note on new evidence *Journal of Policy Analysis and Management, 8*(1), 88-100.

Garen, J. (1988). Compensating wage differentials and the endogeneity of job riskiness. *The Review of Economics and Statistics, 70*(1), 9-16.

Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics, 31*, 93-118.

Gunderson, M., & Hyatt, D. (2001). Workplace risks and wages: Canadian evidence from alternative models. *Canadian Journal of Economics, 34*(2), 377-395.

Hammitt, J. K. (2000). Valuing mortality risk: Theory and practice. *Environmental Science & Technology, 34*(8), 1396-1400.

Hammitt, J. K., & Graham, J. D. (1999). Willingness to pay for health protection: Inadequate sensitivity to probability? *Journal of Risk and Uncertainty 8*(1), 33-62.

Herzog, H. W. J., & M.Schlottmann, A. (1990). Valuing risk in the workplace: Market price, willingness to pay, and the optional provision of safety. *Review of Economics and Statistics, 72*(3), 463-470.

Jones-Lee, M. W. (1974). The value of changes in the probability of death or injury. *Journal of Political Economy, 82*(4), 835-849.

Jones-Lee, M. W., & Loomes, G. (1995). Scale and context effects in the valuation of transport safety. *Journal of Risk and Uncertainty, 11*, 183-203.

Knestaut, A. T. (1997). Fatalities and injuries among truck and taxicab drivers. *Compensation and Working Conditions, Fall 1997*, 55-60.

Kniesner, T. J., Viscusi, W. K., Woock, C., & Ziliak, J. P. (2005). How unobservable productivity biases the value of a statistical Life. *Harvard Law School Discussion paper, 524*.

Kochi, I., Hubbell, B., & Kramer, R. (2006). An empirical Bayes approach to combining and comparing estimates of the value of a statistical life for environmental policy analysis. *Environmental and Resource Economics, 34*(3), 385-406

Lichtenstein, S., Slovic, P., Fischoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal event. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 551-578.

McConnell, K. E. (2006). Chapter 7: Hedonic wage analysis. In N. E. Bockstael & K. E. McConnell (Eds.), *Environmental and Resource Valuation with Revealed Preferences: A Theoretical Guide to Empirical Models (The Economics of Non-Market Goods and Resources)* Springer.

McDaniels, T. L., Kamlet, M. S., & Fischer, G. W. (1992). Risk perception and the value of safety. *Risk Analysis, 12*(4), 495-503.

Mrozek, J. R., & Taylor, L. O. (2002). What determines the value of life? A meta-analysis. *Journal of Policy Analysis and Management, 21*(2), 253-270.

Olson, C. A. (1981). An analysis of wage differentials received by workers on dangerous jobs. *Journal of Human Resources, 16*(2), 167-185.

OMB. (2003). *Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations; Notice*.

Rosen, S. (1974). Hedonic prices and implicit markets: Product diiferentiation in pure competition. *Journal of Polictical Economy, 82*(1), 34-55.

Savage, I. (1993). An empirical investigation into the effect of psychological perception on the willingness to pay to reduce risk. *Journal of Risk and Uncertainty, 6*, 75-90.

Scotton, C., & Taylor, L. (2006). *Of cab drivers and coal miners: pricing heterogeneous risks in the labor market.*Unpublished manuscript.

Scotton, C. R. (2000). *The Impact of Workplace-Risk Measurement on Estimates of the Value of a Statistical Life.* Georgia State University, Atlanta.

Siebert, W. S., & Wei, X. (1994). Compensating wage differentials for workplace accidents: Evidence for union and nonunion workers in the US. *Journal of Risk and Uncertainty, 9*(1), 61-76.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1980). Facts and fears: understanding perceived risk. In R. C. Schwing & J. Walter A. Albers (Eds.), *Societal Risk Assessment : How Safe is Safe Enough?* (pp. 181-216). New York: Plenum Press.

Smith, R. S. (1974). The feasibility of an 'Injury Tax' approach to occupational safety. *Law and Contemporary Problems, 38*(4), 730-744.

Smith, R. S. (1976). *The occupational safety and health act: Its goals and achievements* Washington D.C.: American Enterprise Institute.

Subramanian, U., & Cropper, M. (2000). Public choices between life saving programs: The tradeoff between qualitative factors and lives saved. *Journal of Risk and Uncertainty, 21*(1), 117-149.

Thaler, R., & Rosen, S. (1976). The value of saving a life: Evidence from the labor market. In N. E. Terleckyj (Ed.), *Household Production and Consumption*. New York: Columbia University Press.

USEPA. (1997). *The Benefits and Costs of the Clean Air Act, 1970 to1990*: USEPA.

USEPA. (2004). *Value of Statistical Life Analysis and Environmental Policy: A White Paper*. Washington D.C.: USEPA.

USEPA. (2005). *Regulatory Impact Analysis for the Final Clean Air Interstate Rule*.

USEPA, S. A. B. (2000). *An SAB report on EPA's White Paper: Valuing the Benefits of Fatal Cancer Risk Reduction* Washington D.C.: USEPA.

Viscusi, W. K. (1992). *Fatal Tradeoffs: Public and Private Responsibilities for Risk*. New York: Oxford University Press.

Viscusi, W. K. (2004). The value of life: Estimates with risks by occupation and industry. *Economic Inquiry, 42*(1), 29-48.

Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty, 27*(1), 5-76.

Wooldridge, J. M. (1999). *Introductroy Econometrics: A Modern Approach*: South-Western College Publishing, a devision of Thomson Learning.

Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*: The MIT Press.

Vita

Ikuho Kochi was born in Hiroshima, Japan in 1975 as the second child of Kunitoshi and Takako Kochi. She graduated from Ritsumeikan University, in Kyoto, Japan with the law degree in 1997. She studied environmental cooperation among east and south-east Asian countries in Graduate School for International Development and Cooperation at Hiroshima University, and earned a Master in Science degree in 1999. She studied environmental management at Nicholas School of the Environment and Earth Sciences at Duke University, and earned a Master of Environmental Management degree in 2001. After graduating from Duke University, she continued work on her research project with Dr. Randall Kramer and Dr. Bryan Hubbell as a visiting researcher at Duke University until summer 2002. She entered the Ph.D. in Economics program at Andrew Young School of Policy Studies in 2002 and earned her Ph.D. degree in May 2007. While in the Ph.D. program, she was awarded a Georgia State University Dissertation Grant, a Travel Grant from the International Foundation for Research in Experimental Economics, and the Phi Beta Delta International Scholars Award. Her research appears in peer-reviewed journals such as *Environmental and Resource Economics, Environmental Economics and Policy Studies, Journal of International Development and Cooperation,* and *Journal of Environmental Science.* She has presented her work in various academic conferences and seminars at governmental organizations in Japan, the United States and China. She is currently working for Colorado State University and the U.S. Forest Service Rocky Mountain Research Station as a Post-Doctoral researcher. Ikuho's permanent address is 4-13-14 Kuchita Asakitaku Hiroshima 730-1734 JAPAN.